AD-A046 353 CALIFORNIA UNIV LOS ANGELES SCHOOL OF ENGINEERING A--ETC F/G 17/2
INTEGRATED RANDOM-ACCESS RESERVATION SCHEMES FOR MULTI-ACCESS C--ETC(U)
JUL 77 I RUBIN N00014-75-C-0609

UNCLASSIFIED UCLA-ENG-7752 NL

1 OF 1
ADA
046353

END
DATE
FILMED
12-77
DDC

Research Sponsored by
Office of Naval Research
under Contract N00014-75-C-0609
and the
National Science Foundation
under Grant ENG 75-03224

UCLA-ENG-7752
JULY 1977

INTEGRATED RANDOM-ACCESS RESERVATION SCHEMES FOR MULTI-ACCESS COMMUNICATION CHANNELS

I. RUBIN

UCLA • SCHOOL OF ENGINEERING AND APPLIED SCIENCE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: UCLA-ENG-7752 (14)	2. GOVT ACCESSION NO.
	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): Integrated Random-Access Reservation Schemes for Multi-Access Communication Channels (6)	5. TYPE OF REPORT & PERIOD COVERED
	6. PERFORMING ORG. REPORT NUMBER: UCLA-ENG-7752
7. AUTHOR(s): Izhak Rubin (10)	8. CONTRACT OR GRANT NUMBER(s): N00014-75-C-0609, NSF-ENG-75-03224 (15)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: University of California, Los Angeles, CA 90024	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR042-291
11. CONTROLLING OFFICE NAME AND ADDRESS: Statistics and Probability Program, Office of Naval Research, Arlington, Virginia 22217	12. REPORT DATE: July 1977 (11)
	13. NUMBER OF PAGES: 71
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): (12) 77 p.	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Reproduction in whole or in part is permitted for any purpose of the United States Government.

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Communication Networks
Data Networks
Computer Communications
Satellite Communications
Multiple-Access Techniques
Random-Access Schemes
Packet-Switching Networks

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Integrated Random-Access Reservation (IRAR) access-control schemes, for multi-access communication channels, are presented and studied. Under an IRAR scheme, newly arrived packets can be designated for reserved or random-access transmissions. In the latter case, if a collision occurs, each colliding packet is assigned for transmission by reservation, rather than attempt another random-access transmission.

OVER

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

404 637

Block 20 (Continued)

The performance of the IRAR schemes ~~under consideration~~ is expressed in terms of the associated channel delay-throughput curves. Single packet messages are assumed. Analytical expressions are derived for the limiting average packet delay for a broadcast multi-access channel, under various IRAR schemes. The performance curves under IRAR access-control schemes are compared with those resulting when pure reservation and pure random-access procedures are used. ~~We show~~ that for channels with low propagation delays (such as terrestrial radio or line channels), an IRAR scheme yields an excellent performance curve over the whole throughput range. For channels with high propagation delays (such as a satellite communication channel), an IRAR scheme is shown to yield delay-throughput performance characteristics superior to those obtained under pure reservation and random-access schemes, for medium and low network throughput values. At higher throughput values, an efficiently designed integrated access control scheme will be appropriately switched from an IRAR mode into a reservation procedure.

UCLA-ENG-7752
JULY 1977

INTEGRATED RANDOM-ACCESS RESERVATION SCHEMES FOR MULTI-ACCESS COMMUNICATION CHANNELS

Izhak Rubin

Research Sponsored by
Office of Naval Research
under Contract N00014-75-C-0609
and the
National Science Foundation
under Grant ENG75-03224

Department of System Science
School of Engineering and Applied Science
University of California
Los Angeles, California 90024

Abstract

Integrated Random-Access Reservation (IRAR) access-control schemes, for multi-access communication channels, are presented and studied. Under an IRAR scheme, newly arrived packets can be designated for reserved or random-access transmissions. In the latter case, if a collision occurs, each colliding packet is assigned for transmission by reservation, rather than attempt another random-access transmission.

The performance of the IRAR schemes under consideration is expressed in terms of the associated channel delay-throughput curves. Single packet messages are assumed. Analytical expressions are derived for the limiting average packet delay for a broadcast multi-access channel, under various IRAR schemes. The performance curves under IRAR access-control schemes are compared with those resulting when pure reservation and pure random-access procedures are used. We show that for channels with low propagation delays (such as terrestrial radio or line channels), an IRAR scheme yields an excellent performance curve over the whole throughput range. For channels with high propagation delays (such as a satellite communication channel), an IRAR scheme is shown to yield delay-throughput performance characteristics superior to those obtained under pure reservation and random-access schemes, for medium and low network throughput values. At higher throughput values, an efficiently designed integrated access control scheme will be appropriately switched from an IRAR mode into a reservation procedure.

I. Introduction and System Description

We consider a multi-access communication channel of capacity C bps serving a network of terminals. We assume information transmitted through the channel is broadcasted to all network terminals, so that each terminal is able to listen to any other terminal in the network. Subsequently, decentralized access-control procedures can be employed. For this purpose, it is further assumed that each terminal records in its own queueing table the relevant state of the channel utilization process. Non-broadcast channels with centralized control functions are readily designed and analyzed using direct modifications of the techniques and results presented here. A satellite communication channel and a terrestrial radio channel serve as examples of multi-access broadcast radio communication channels. Such channels utilize a repeater (such as a satellite transponder or a radio relay station) to allow each terminal in the network to communicate with any other terminal. We assume the corresponding channel uplink and downlink (constituting the terminal-repeater and repeater-terminal links, respectively) to possess disjoint frequency bands, so that users need to contend only for the use of the uplink channel. A communication line in a line computer communication network serves as another example. Terminals, wishing to share the use of this line to transmit their information to an appropriate destination node, can be located at one end of this line or distributed (multi-dropped) geographically along it.

We assume a synchronized structure. Time (referenced to repeater's time) is thus divided into fixed-length durations of τ sec. each, called slots. Terminals will start message transmissions only at times coinciding with the starting times of the synchronized time slots. The channel is characterized by a propagation delay of $R\tau$ sec., or R slots. Propagation delays are short for regular terrestrial radio or line channels and are longer for satellite networks and long-distance terrestrial channels.

Considering a network of M terminals, new messages arrive at the i-th terminal according to a Poisson stream of intensity λ_i [mess./slot], i=1,2,...,M. The overall network message arrival stream is thus a Poisson point process with intensity $\lambda = \sum_{i=1}^{M} \lambda_i$ [mess./slot]. Messages are assumed here to be composed of a single fixed length packet, containing μ^{-1} [bits/packet]. The packet transmission-time through the channel is thus $(\mu C)^{-1}$ [sec./packet]. We set, as in [1] - [2], the slot duration to equal the packet transmission time, so that $\tau = (\mu C)^{-1}$.

A terminal will try to gain channel access for its packets, immediately upon their arrival, following the protocol governing the network underlying access-control discipline. Considering terminals which generate bursty (low duty-cycle, high peak-to-average traffic intensity ratio) message streams, an appropriate access-control discipline needs to be chosen so that the bandwidth of the channel is utilized efficiently and acceptable message queueing delays are obtained. Such efficient decentralized message-switching access-control schemes have been recently developed and studied in [1] - [2]. (See [1] - [2]. for references to other related studies of access-control schemes.)

Reservation access-control schemes have been studied in [1]. These schemes have shown to yield excellent delay-throughput performance characteristics at medium and high network traffic intensity values. Messages containing a random number of packets have been incorporated in these studies. In analizing reservation schemes, the notion of a contention-free reservation slot has been utilized. Within such a slot, all terminals are assumed to be able to successfully transmit their reservation packets (or mini-packets, see [1]). In particular, we mention here the following reservation schemes studied in [1]. A Fixed Reservation Access Control (FRAC) scheme utilizes a fixed pattern consisting of a single pre-determined reservation slot followed by a number (N) of service slots. For each arriving message, the terminal thus sends its reservation packet in the reservation slot nearest to the arrival, and is subsequently assigned the appropriate service slots for the transmission of the message. When the number N is dynamically changed to appropriately match the underlying network traffic intensity value λ, a Dynamic Fixed Reservation Access Control (DFRAC) scheme results. The latter has been shown in [1] to yield low average message delay values over the whole practical range of medium to high network traffic intensity values.

A reservation scheme which automatically changes its structure to accomodate network traffic fluctuations has also been presented and studied in [1]. This is the Aynchronous Reservation Demand-Assignment ARDA I scheme. This scheme declares each idle, unreserved, slot as a reservation slot. Other slots are used for packet transmissions. In this way, the frequency of reservation slots is set automatically in accordance with the network traffic value. ARDA I schemes have shown to yield excellent delay-throughput characteristics (somewhat superior to those of a DFRAC scheme) for channels with low propagation delay values, R = 0,1. For channels with higher propagation delay values, it has been shown in [1] that the DFRAC scheme exhibits better delay-throughput performance characteristics. (An ARDA II scheme has been developed there to yield automatic dynamic adaptation to traffic intensity fluctuations while exhibiting a delay-throughput performance similar to that of a DFRAC scheme). The performance curves of the latter schemes will be compared with those of the integrated schemes introduced and studied in this paper.

For low network traffic intensity values, when single-packet bursty terminal message streams are considered, a better delay-throughput performance, involving a much less sophisticated distributed access-control procedure, can be achieved by a random-access mechanism. Terminals then can use the channel at any time to transmit a newly arrived packet. If, however, two or more packets collide, the involved messages are retransmitted following an appropriate random retransmission delay policy. A Group Random-Access (GRA) procedure has been introduced and studied in [2]. Under this scheme, a family of network terminals is allowed to use only a specified (periodic) pattern of channel time epochs (groups), on a random-access basis. (At other times, channels utilization is governed by possibly other access-control strategies, and can be dedicated to various other classes of terminals demanding a different type of service.) Packets colliding within a certain group of slots will thus retransmit within the next prescribed group of slots. As for the slotted ALOHA random-access procedure, the GRA scheme has been shown in [2] to have traffic capacity of $e^{-1} = 0.368$ [packets/slot] and be inherently unstable. To stabilize the GRA scheme, a dynamic feedback optimal control procedure have been derived in [2]. The latter needs to reject packets from the system, at certain times, yielding the minimal average packet delay attainable at an acceptable prescribed probability of packet rejection. The controlled scheme subsequently yields very low average packet delay values at low enough network throughput values. In particular, we note that at very low network throughput values (as $\lambda \to 0$), the random-access scheme can yield a reduction of up to R+1 slots in packet delay, over that of a reservation scheme, due to the related saving of the propagation and transmission delay involved with the broadcast transmission of a reservation packet. Furthermore, the simple distributed control mechanism involved with a random-access discipline, can result in significant savings in hardware requirements and reductions in protocol and system complexities.

It is the purpose of this paper to integrate the distinct advantages of reservation and random-access disciplines, and introduce and study access-control schemes which combine both random-access and reservation operations.

Such Integrated Random-Access Reservation (IRAR) schemes are expected to yield excellent delay-throughput characteristics over a very wide range of network traffic intensity (throughput) values. IRAR schemes will also be stable over the whole range of allowable traffic intensity values ($0 \le \lambda < 1$), due to the availability of the reservation operation, eliminating thus the need for a complexed stabilizing control mechanism for the random access operation.

The basic governing principle involved in the operation of the various IRAR schemes under consideration in this paper, is described as follows. A newly arriving packet is allowed (many times) to be transmitted in certain time slots on a random-access basis. If a subsequent collision has occurred, this packet is however not retransmitted again on a random-access basis (as is the case under a random-access discipline) but is instructed to use a reservation procedure. The resulting IRAR schemes thus switch from a majority of random-access transmissions at low network throughput values to a majority of reserved transmissions at higher throughput values.

The protocols of the basic IRAR schemes are presented in Section II. For these schemes, the delay-throughput performance characteristics are derived in Section III, assuming channels with low propagation delay values, $R = 0$, as for terrestrial radio or line communication networks. For higher propagation delay values, the delay-throughput performance of IRAR schemes, is presented in Section IV. The performance curves of the various IRAR schemes are subsequently compared with those of pure random-access schemes. The performance advantage of IRAR schemes, in many situations, is consequently concluded. In addition to demonstrating the performance characteristics of the basic IRAR schemes, our studies here will also serve as the basic background for extensions to even more sophisticated integrated access-control disciplines, for communication networks involving a multitude of different types of messages.

II. Protocols for the Basic IRAR Schemes

To specify the protocol of an IRAR scheme, we make use of the notions of a reservation slot and a random-access slot. A reservation slot is defined to be a slot which is dedicated, as specified by the protocol of the underlying access-control discipline, for reservations. Furthermore, we assume that within a reservation slot every network terminal can broadcast its reservation packet (or mini-packet), anouncing its requirements for channel time (and priorities, if desired), in a contention free manner (see [1] for further details). We note that if a discipline requiring access contention is employed to govern reservation transmissions within reservation slots (such as a slotted ALOHA procedure, so that a GRA scheme results), the resulting extra reservation delay factors can be added to the delay formulas presented here (with λ now recognized as the throughput rate of the stream of reserved messages, assuming the latter is modelled as a Poisson point process).

A random-access (RA) slot is declared to be an unreserved slot during which terminals are allowed to transmit their packets on a random-access (unreserved) basis. Clearly, a successful packet transmission within a RA slot can occur if and only if a single terminal transmits a packet within this slot. If two or more terminals transmit packets within the same RA slot, their packets will collide and subsequently be recognized by the network terminals (R slots following their transmission) as unsuccessful random-access transmissions. In the latter case, the corresponding terminals, using copies of these packets (alway kept in the terminal buffer till packets are acknowledges to be successfully transmitted), will try to gain channel access again. A slot during which packet collisions occur is called a collision slot. In the IRAR schemes studied here, a packet is allowed to try no more than a single random-access transmission. A colliding packet will be instructed by the IRAR scheme to use a reservation procedure to gain access, and send its reservation packet within an appropriately scheduled reservation slot. Slots reserved for packet transmissions are called service slots.

It is also possible to set the allowable number of packet random-access transmissions to be equal to a fixed number L, L ≥ 1. Following L unsuccessful transmissions, a packet is then instructed to use a reservation procedure. However, we will note that an IRAR scheme with L = 1, essentially attains the delay-throughput performance of a pure random-access scheme (such as a GRA or slotted ALOHA scheme) over the appropriate range of network throughput values. Therefore, an IRAR procedure employing a value of L > 1 is not expected to yield a scheme with improved performance characteristics, as actually observed through various simulation studies.

It is further interesting to note that the limiting case of L = 0 can be viewed as corresponding to a situation where an arriving packet sends immediately a reservation packet, at the first available idle slot, not using any RA transmission trials at all. The latter procedure, however, constitutes the ARDA I reservation scheme mentioned in Section I.

The first basic IRAR scheme, denoted as IRAR I, is administered by the following protocol.

<u>Protocol for IRAR I Scheme</u>:

1. Any slot which has not been reserved as a service or reservation slot, is declared to be a random-access (RA) slot.

2. The (first unreserved) slot following the instant at which a collision is recognized by the network terminals (i.e., R slots following any collision slot) is established as a reservation slot.

3. Service slots are established (at unreserved slots) for reserved messages immediately following the reception of the corresponding reservation packet

(i.e., R slots after the corresponding reservation slot), in accordance with the underlying service ordering discipline.

4. A newly arriving message transmits its (information carrying) packet in the first available RA slot. If collision is subsequently recognized (R slots latter), reservation is made for this packet in the first available reservation slot.

We note that instructions 1, 2 and 3 of the IRAR I protocol describe the establishment of random-access, reservation and service slots, respectively. Each terminal, using its queueing table, thus recognizes the identity of the underlying slots, and subsequently employs instruction 4 to manage its transmissions.

To determine the order of service among packets which have made reservations within the same reservation slot, any service ordering discipline (such as random ordering, first-come first-served or any priority ordering procedure) can be used. We however assume here that packets which have made reservations at an earlier reservation slot are served before those which make reservations at a latter reservation slot.

To improve the delay-throughput performance of an IRAR I scheme at medium and high network traffic flow values, we modify this scheme as follows. If a newly arriving message recognizes a reservation slot, prior to the first available RA slot, we allow it to make immediately its reservation within the latter reservation slot, rather than try first RA transmission. The resulting access procedure,denoted as IRAR II scheme, is thus governed by the following protocol.

Protocol for IRAR II Scheme: Instructions 1, 2 and 3 for IRAR II are identical to instructions 1, 2 and 3, respectively, for IRAR I. Instruction 4 for IRAR II is given as follows. (4) Upon its arrival, a newly arriving message checks whether there exists any reservation slot preceding the first available RA slot.

If the latter is the case, the message sends a reservation packet in the latter reservation slot. Otherwise, the information packet is transmitted in the first available RA slot. Then, in the latter situation, if collision is subsequently recognized, reservation is made for this packet in the first available reservation slot. |

A further improvement in the delay-throughput curve of an IRAR scheme, at medium and high network throughput values, can be achieved as follows. We note that at such traffic intensity values, service periods (being groups composed of successive service slots) can be relatively long. Subsequently, a high probability of packet collisions will exist within the RA slot following a service period. To avoid such collisions, an IRAR III scheme declares the latter slot to be a reservation slot, assuming thus the following protocol.

Protocol for IRAR III Scheme: Instructions 1, 3 and 4 for IRAR III are identical to instruction 1, 3 and 4, respectively, for IRAR II. Instruction 2 for IRAR III is given as follows. (2) The (first unreserved) slot following the instant at which a collission is recognized by the network terminals, is established as a reservation slot. Also, any (unreserved) slot following a service slot is declared as a reservation slot. |

The delay-throughput performance characteristics of IRAR I, II and III schemes, as well as other related IRAR schemes, are derived in Section III for short propagation delay values ($R = 0$) and in Section IV for longer propagation delay values ($R \geq 1$).

III. Performance Analysis of IRAR Schemes for Channels with R = 0

The Performance Measures

The performance of the IRAR schemes under consideration here will be assessed in terms of their delay-throughput curves. The latter relate the (steady-state) average packet delay D in the network to the network throughput s. The message delay constitutes of the following components. The waiting-time of the n-th packet in the system is denoted as W_n, $n \geq 1$. This waiting-time is expressed in terms of numbers of slots and is measured from the start of the slot following the packet's arrival at its terminal to the instant this packet is successfully transmitted. Note that we do not include an average delay of 1/2 slot which accounts for the average time elapsing between the actual packet arrival (following the continuous-time Poisson stream) and the start of the next slot. The steady-state average packet waiting-time function $\bar{W}$ is given by the limit (when it exists)

$$\bar{W} = \lim_{n \to \infty} N^{-1} E\left\{\sum_{n=1}^{N} W_n\right\} . \tag{3.1}$$

The overall steady-state average message delay D (in slots) is thus given by

$$D = \bar{W} + R + 1 . \tag{3.2}$$

Equation (3.2) includes R slots and single slot terms to account for propagation delay and transmission time, respectively, associated with a successful packet transmission.

Denoting by S_i the number of successful packet transmissions in the i-th slot, $S_i = 0, 1$, $i \geq 1$, the channel throughput is given by

$$s = \lim_{N \to \infty} N^{-1} E\left\{\sum_{i=1}^{N} S_i\right\} , \tag{3.3}$$

yielding the channel output rate (i.e., the limiting average number of successful packet transmissions per slot).

For all IRAR schemes considered here, it is readily verified that the underlying Markov chain describing the channel state process is positive-recurrent if and only if $\lambda < 1$. Assuming thus henceforth that $\lambda < 1$, and since no packet rejections are used in IRAR schemes, we have

$$s = \lambda \ , \tag{3.4}$$

so that (at steady-state) the channel throughput is equal to the overall network traffic intensity value.

The Analytical Technique for Evaluating the Packet Delay

The analytical procedure presented in [1] is used also here to compute the packet delay. It is briefly summarized as follows. The channel state process is described by a vector Markov chain $\{\underline{X}_n, n \geq 0\}$. Associated with any sample function of the channel state process, appropriate time periods are defined and identified, under each IRAR scheme. Random vector $\underline{X}_n$ thus describes the state of the n-th period. (A typical time period will include a collision slot and the associated reservation and service slots.) We then set $N(\underline{X}_n, \underline{X}_{n+1})$ and $W(\underline{X}_n, \underline{X}_{n+1})$ as the average (given $\underline{X}_n, \underline{X}_{n+1}$) values of the number of messages served and the sum of waiting-times of these messages, respectively, during the (n+1)-st time period associated with state $\underline{X}_{n+1}$. For our schemes, the latter two functions are time-homogeneous and depend only on $(\underline{X}_n, \underline{X}_{n+1})$, for each $n \geq 0$. We further have, as $M \to \infty$, for $\lambda > 0$,

$$\sum_{n=1}^{M} N(\underline{X}_n, \underline{X}_{n+1}) \to \infty \ , \quad \text{w.p.1.} \tag{3.5}$$

Subsequently, we can write

$$\lim_{N \to \infty} N^{-1} \sum_{n=1}^{N} W_n = \lim_{M \to \infty} \frac{\sum_{n=0}^{M} W(\underline{X}_n, \underline{X}_{n+1})}{\sum_{n=0}^{M} N(\underline{X}_n, \underline{X}_{n+1})} \ , \tag{3.6}$$

with probability one. We now apply a Markov ratio limit theorem (see [3], p. 91, Theorem 1, and [1]) to the vector Markov chain $\{\underline{Y}_n, n \geq 1\}$, where $\underline{Y}_n = \{\underline{X}_n, \underline{X}_{n+1}\}$. For $\lambda < 1$ and any IRAR scheme, the latter chain is an irreducible positive-recurrent Markov chain with the stationary distribution $\{\pi(\underline{i},\underline{j})\}$. We then conclude that

$$\bar{W} = \frac{E[W(\underline{X}_n, \underline{X}_{n+1})]}{E[N(\underline{X}_n, \underline{X}_{n+1})]} , \tag{3.7a}$$

where

$$E[W(\underline{X}_n, \underline{X}_{n+1})] = \sum_{\underline{i},\underline{j}} W(\underline{i},\underline{j})\pi(\underline{i},\underline{j}) , \tag{3.7b}$$

$$E[N(\underline{X}_n, \underline{X}_{n+1})] = \sum_{\underline{i},\underline{j}} N(\underline{i},\underline{j})\pi(\underline{i},\underline{j}) . \tag{3.7c}$$

Thus, to evaluate the limiting average packet waiting-time $\bar{W}$, we need to obtain functions $W(\cdot)$ and $N(\cdot)$, and derive the steady-state joint probabilities $\{\pi(\underline{i},\underline{j})\}$. The latter can be expressed, for each $\underline{i}$, $\underline{j}$, as

$$\pi(\underline{i},\underline{j}) = u(\underline{i})P(\underline{i},\underline{j}) , \tag{3.8}$$

where $\{P(\underline{i},\underline{j})\}$ and $\{u(\underline{i})\}$ denote the transition probability function and the stationary distribution, respectively, of Markov chain $\{\underline{X}_n\}$. For the schemes under consideration, we will note that it is not necessary to fully have $\{u(\underline{i})\}$. The stationary probabilities of only a limited number of states (state 0 alone, or states 0 and 1, usually) will be required. We note that the same procedure is readily extended to evaluate any other limiting moment of the message waiting-time.

Performance Analysis for Channels with R = 0.

The IRAR I Scheme

In this section we will consider channels with low propagation delay value, setting $R = 0$. The analytical procedure presented above will be used to derive the delay-throughput performance curves under IRAR I, II and III schemes.

We consider first an IRAR I scheme for a multi-access channel with R = 0. For this situation, a sample function of the process describing the evaluation of the channel reservation and service periods is shown in Fig. 3.1. We note that the channel state process is composed of a series of basic time periods. Such a period is set so that its first slot is always a RA slot. A period constitutes of a succession of a single collision slot, a single reservation slot and a number of service slots (a service period), provided that the first slot is a collision slot. Otherwise, the first (RA) slot contains no collissions, and we take it to constitute the whole corresponding time period.

Using the above mentioned definition of a period, we let $\underline{X}_n$ denote the state vector associated with the n-th period, so that $\{\underline{X}_n, n \geq 0\}$ is the underlying state sequence. We let $N_n = X_n$ denote the number of packets successfully transmitted within the n-th period. Clearly, we can use here $\underline{X}_n = \{X_n\}$, so that the state sequence is represented by the scalar chain $\{X_n, n \geq 0\}$. For a single slot n-th period, we have $X_n = 0$ if no new arrivals occur, or $X_n = 1$ if a single successful RA transmission occurs. For a longer n-th period, X_n denotes the number of colliding, reserved and served packets within this period.

We readily note that $\{X_n, n \geq 0\}$ is a Markov chain, with a state-space composed of the nonnegative integers $\mathcal{J}$. Its transition probability function $\{P(i,j)\}$ is described by the transion diagram of Fig. 3.2. Thus, we have

$$P(i,j) = P(X_{n+1}=j \mid X_n=i) = \begin{cases} [\lambda(i+2)]^j (j!)^{-1} \exp(-\lambda(i+2)), & \text{if } i \geq 2, \\ \lambda^j (j!)^{-1} \exp(-\lambda), & \text{if } i = 0,1, \end{cases} \tag{3.9}$$

for $j \geq 0$, $n \geq 0$. We note, as indicated in Eq. (3.9) or Fig. 3.2 that the number of RA transmissions at the first slot of each group is equal to the number of new packet arrivals in the previous group. If the latter is i slots long, $i+2$ slots are available for these new arrivals if $i \geq 2$, since the collision and reservation slots are included. All RA transmissions are served within the group. For $\lambda < 1$, one readily verifies that $\{X_n, n \geq 0\}$ is positive-recurrent, having a stationary *distribution* $\{u_i, i \geq 0\}$, where

$$u_i = \lim_{n \to \infty} P(X_n = i) , \quad i \geq 0 . \tag{3.10}$$

Distribution $\{u_i\}$ is obtained as the unique solution to

$$u_j = \sum_{i=0}^{\infty} u_i P(i,j), \ j \geq 0 , \quad \sum_{j=0}^{\infty} u_j = 1 . \tag{3.11}$$

It is, however, not necessary for our purposes to solve (3.11) for the steady-state distribution. It will be shown that only u_o, u_1 are needed. The latter probabilities can be obtained from (3.9) - (3.11), or more simply by a simulation of $\{X_n\}$, following the flow diagram of Fig. 3.2. In the latter case, we set (using a sample mean estimate)

$$u_i(M) = M^{-1} \sum_{n=0}^{M-1} I(X_n = i), \tag{3.12}$$

with $u_i = \lim_{M \to \infty} u_i(M)$, where $I(A)$ is the indicator function associated with event A, so that $I(A) = 1$ if A has occurred and $I(A) = 0$, otherwise. Due to the simplicity of the transition scheme of Fig. 3.2, it is easy and simple to use such a simulation, or recursive computation procedure to obtain u_o and u_1. We thus assume henceforth that u_o and u_1 are obtained in this manner (also been used in the actual computations).

To evaluate the average packet waiting-time $\bar{W}$, we now need functions $N(\cdot)$. The number of packets served during the (n+1)-st group is clearly equal to X_{n+1}. We thus have

$$N(X_n, X_{n+1}) = X_{n+1} . \tag{3.13}$$

The sum of the waiting-times (in number of slots) of packets served during the (n+1)-st period, for an IRAR I scheme, is given by

$$W(X_n, X_{n+1}) = I(X_n \geq 2)\cdot[\frac{1}{2} X_{n+1}(X_n+1)]$$
$$+ I(X_{n+1} \geq 2)\cdot[2X_{n+1} + \frac{1}{2} X_{n+1}(X_{n+1} - 1)] \ . \qquad (3.14)$$

The first term on the RHS of (3.14) expresses the overall (averaged, w.r.t. to their Poisson arrival times, given X_n, X_{n+1}) waiting time of packets from the time of their arrival within the n-th period to the time of their RA transmission at the first slot of the (n+1)-st period. For X_n=0,1 , this term should be set to 0 , as ensured by the multiplication by $I(X_n \geq 2)$. The second term of (3.14) yields the overall waiting-time of packets within the (n+1)-st period, from their RA transmission to their admission into service. This term is set to zero if X_{n+1}=0 or 1 . Otherwise, the first component, $2X_{n+1}$, accounts for the overall delay experienced due to transmissions over a collision and a reservation slot. The second component yields the averaged (given X_{n+1}) waiting-time of a packet, for admission into service, following its acknowledged reservation, given that X_{n+1} packets will be scheduled for service (at random order). (Note that since only the average waiting-time is computed, any "conserving" service ordering discipline will yield the same overall waiting-time value, and is therefore acceptable for our computations.)

Incorporating (3.9)-(3.14), the delay-throughput performance curve of an IRAR I scheme, when R=0 , can be evaluated using Eqs.(3.7). For that purpose,we first derive expressions for the first two (steady-state) moments of $\{X_n\}$, in terms of u_0, u_1 . The latter computation follows from Eq.(3.9), or the diagram of Fig. 3.2 . The latter is represented by the following recursive state relationship,

$$X_{n+1} = I(X_n \geq 2)Z_n(X_n + 2) + I(X_n \leq 1)Z_n(1) \ , \tag{3.15}$$

where $Z_n(i)$ is a sequence of i.i.d. random-variables governed by a Poisson distribution with intensity λi, $i \geq 1$. The limiting moments, for $\lambda < 1$,

$$\bar{X} = \lim_{n \to \infty} E(X_n) = \sum_i i u_i \ , \tag{3.16a}$$

$$\bar{X}^2 = \lim_{n \to \infty} E(X_n^{\ 2}) = \sum_i i^2 u_i \ , \tag{3.16b}$$

are subsequently given as follows.

<u>Lemma 3.1</u> For $\lambda < 1$, $R = 0$, under and IRAR I scheme, limiting moments $\bar{X}$ and $\bar{X}^2$ exist and are given by

$$\bar{X} = \lambda(1 - \lambda)^{-1}[2 - u_0 - 2u_1] \ , \tag{3.17a}$$

$$\bar{X}^2 = (1 - \lambda^2)^{-1}[\bar{X}(1 + 4\lambda^2) + 4\lambda^2 - 5\lambda^2 u_1 - 3\lambda^2(u_0 + u_1)], \tag{3.17b}$$

where u_0, u_1 are given by (3.10).

<u>Proof</u>. Eq.(3.17a) follows by taking expectations on both sides of (3.15), noting that at equilibrium,

$$E\{I(X_n \geq 2)Z(X_n + 2)\} = E\{\lambda(X_n + 2)I(X_n \geq 2)\}$$

$$= \lambda(\bar{X} - u_1) + 2\lambda(1 - u_0 - u_1) \ ;$$

$$E\{I(X_n \leq 1)Z(1)\} = \lambda(u_0 + u_1) \ ;$$

and setting $\lim E(X_n) = \lim E(X_{n+1}) = \bar{X}$.

Eq.(3.17b) follows by squaring both sides of (3.15), taking expectations, and noting that at equilibrium,

$$E\{I^2(X_n \geq 2)Z^2(X_n + 2)\} = E\{I(X_n \geq 2)[\lambda(X_n + 2) + \lambda^2(X_n + 2)^2]\ ;$$

$$E\{I^2(X_n \leq 1)Z^2(1)\} = E\{I(X_n \leq 1)(\lambda^2 + \lambda)\} = (u_0 + u_1)(\lambda^2 + \lambda)\ ,$$

and setting $\lim E(X_{n+1}{}^2) = \lim E(X_n{}^2) = \bar{X}^2$

Q.E.D.

The delay-throughput performance curve is now obtained in Theorem 3.1.

Theorem 3.1 For a multi-access channel with $R = 0$, under an IRAR I scheme, for a packet arrival intensity of $\lambda < 1$, the limiting average packet delay D is given by

$$D = \bar{W} + 1\ , \tag{3.18}$$

where

$$\bar{W} = 1.5(1 + \lambda) + \frac{1}{2}(1 + \lambda)\bar{X}^2(\bar{X})^{-1} + (\bar{X})^{-1}[\lambda(1 - u_0 - u_1) - 2(1 + \lambda)u_1]\ , \tag{3.19}$$

where $\bar{X}$ and $\bar{X}^2$ are given by Eqs.(3.17). For $\lambda \geq 1$, D is unbounded.

Proof. Eq.(3.19) is obtained using Eqs.(3.7).

By (3.13), we obtain

$$E[N(X_n, X_{n+1})] = \bar{X}\ .$$

By (3.14), we have

$$E[W(X_n, X_{n+1})] = \frac{1}{2}\lambda(\bar{X}^2 - u_1) + \lambda(\bar{X} - u_1) + \frac{1}{2}\lambda(\bar{X} - u_1) + \lambda(1 - u_0 - u_1) + \frac{3}{2}(\bar{X} - u_1) + \frac{1}{2}(\bar{X}^2 - u_1)\ ;$$

noting that, at equilibrium (using relation (3.15)),

$$E\{X_n X_{n+1} I(X_n \geq 2)\} = E\{X_n I(X_n \geq 2) E[X_{n+1} | X_n \geq 2]\} = E\{X_n I(X_n \geq 2)\lambda(X_n + 2)\}$$

$$= \lambda E\{X_n^2 I(X_n \geq 2)\} + 2\lambda E\{X_n I(X_n \geq 2)\}$$

$$= \lambda(\bar{X}^2 - u_1) + 2\lambda(\bar{X} - u_1) \ ;$$

and that

$$E\{X_{n+1} I(X_n \geq 2)\} = \lambda E\{(X_n + 2) I(X_n \geq 2)\}$$

$$= \lambda E\{X_n I(X_n \geq 2)\} + 2\lambda E\{I(X_n \geq 2)\}$$

$$= \lambda(\bar{X} - u_1) + 2\lambda(1 - u_0 - u_1) \ .$$

Eq.(3.18) then follows from (3.2). Since the channel process approaches a pure reservation process as $\lambda \uparrow 1$, $\lambda < 1$ as the NASC for its positive-recurrence follows as for a reservation scheme in [1], or directly by (3.15).

Q.E.D.

Note, as a check of (3.19), that as $\lambda \to 0$, we have $u_0 \to 1$, $u_1 \to 0$, $\lambda(\bar{X})^{-1} \to 1$, $\bar{X}^2(\bar{X})^{-1} \to 1$, $u_1\lambda^{-1} \to 1$, so that $\bar{W} \to \frac{3}{2} + \frac{1}{2} - u_1(\bar{X})^{-1} \to 0$, as expected.

Eqs.(3.17)-(3.19) thus yield the delay-throughput curves for $R = 0$, under an IRAR I scheme, with probabilities u_0, u_1 as parameters. The latter probabilities are evaluated as indicated above, through (3.12) or (3.11). The resulting delay-throughput curve is shown in Fig. 3.7. It will be compared later with curves obtained under other access-control schemes.

The IRAR II Scheme

We now consider a multi-access channel with R = 0 under an IRAR II scheme. A sample function of the associated channel state process is shown in Fig. 3.3. Time periods are defined as for the IRAR I scheme. Thus, the first slot in any time period is always a RA slot. Considering now the n-th period, the number of transmissions in this RA slot is denoted as N_n. The number of successful transmissions within the n-th period is denoted again as X_n. If $N_n = 0$, or $N_n = 1$, then $X_n = N_n$ and a single slot group containing no transmissions, or a single successful transmission, results. On the other hand, if $N_n \geq 2$, a collission slot results. This slot is followed by a reservation slot. The number of packets making reservations at the latter slot is given by $N_n + I_n$. This includes the N_n colliding packets plus I_n new packets which have arrived during the previous collission slot. Variable I_n represents the new parameter incorporated in IRAR II, when compared with IRAR I. It represents those packets which have arrived during a collission slot and are thus allowed by instruction 4 of the IRAR II protocol to immediately make a reservation at the next slot, which is established as a reservation slot. Under the IRAR I protocol, such newly arriving packets would have to wait for the next RA slot and then execute a RA transmission. The latter will result with a high probability of collision, under medium and high network throughput values. Under lower network traffic intensity values, IRAR I and II clearly exhibit similar performance characteristics. We thus expect an improvement in the performance characteristics when using an IRAR II scheme, over the whole throughput range.

The channel Markov state sequence $\{\underline{X}_n, n \geq 0\}$ is now characterized by the state $\underline{X}_n$, where

$$\underline{X}_n = (N_n, I_n) . \tag{3.20}$$

The state equations follow the flow diagram of Fig. 3.4, and are described as follows. The number of packets successfully transmitted in the (n+1)-st group, X_{n+1}, is given by

$$X_{n+1} = N_{n+1} + I_{n+1}\, I(N_{n+1} \geq 2), \tag{3.21}$$

where

$$P\{N_{n+1}=j \mid N_n, X_n=i\} = \begin{cases} [\lambda(i+1)]^j (j!)^{-1} \exp(-\lambda(i+1)), & \text{if } i \geq 2, \\ \lambda^j (j!)^{-1} \exp(-\lambda), & \text{if } i = 0,1, \end{cases} \tag{3.22}$$

and $\{I_n, n \geq 0\}$ is a sequence of i.i.d. random-variables governed by a Poisson distribution with mean λ,

$$P\{I_{n+1} = j\} = \lambda^j (j!)^{-1} \exp(-\lambda) , \tag{3.23}$$

for $j \geq 0$.

Eq. (3.21) indicates that $X_{n+1} = N_{n+1}$ for a single slot group, while $X_{n+1} = N_{n+1} + I_{n+1}$ (so that the N_{n+1} colliding packets plus the newly arrived I_{n+1} packets are served) for a group containing a colliding first slot. Eq. (3.22) indicates that the number N_{n+1} of RA transmissions made in the (n+1)-st slot is equal to the number of new arrivals during the X_n+1 service plus reservation slots of the n-th group, if $X_n \geq 2$. If $X_n \leq 1$, the n-th period contains a single slot and N_{n+1} contains only the number of new arrivals within this slot. Eq. (3.23) indicates that the number of packets I_{n+1} making reservations without first trying RA transmission, is equal to just those newly arriving within the single colliding slot.

To evaluate the packet average waiting-time we now represent functions $N(\cdot)$ and $W(\cdot)$. For the overall number of packets served during the (n+1)-st group, we have again

$$N(\underline{X}_n, \underline{X}_{n+1}) = X_{n+1} . \tag{3.24}$$

The overall sum of waiting-times for packets served during the (n+1)-st period, is given by

$$W(\underline{X}_n, \underline{X}_{n+1}) = I(X_n \geq 2) \cdot \left[\frac{1}{2} N_{n+1} X_n\right]$$
$$+ I(X_{n+1} \geq 2) \left[N_{n+1} + X_{n+1} + \frac{1}{2} X_{n+1}(X_{n+1} - 1)\right]. \qquad (3.25)$$

Comparing Expressions (3.25) and (3.14), we note that the first expression on the RHS of (3.25) again expresses the overall (averaged) waiting time of those N_{n+1} packets arriving during the last (X_n+1) slots of the n-th group, prior to their RA transmission in the first slot of the (n+1)-st group. Similarly, the second term incorporates delay times involving, provided $X_{n+1} \geq 2$ (noting that $X_{n+1} \geq 2 \Leftrightarrow N_{n+1} \geq 2$): a collision slot for each one of N_{n+1} packets; a reservation slot for each one of X_{n+1} packets; and an average overall waiting-time from reservation to admission into service of $\frac{1}{2} X_{n+1}(X_{n+1}-1)$ slots.

We note that $\{X_n, n \geq 0\}$ is itself a Markov chain, having a transition probability function $\{P(i,j)\}$ given by (3.21) - (3.23), a stationary distribution $\{u_i, i \geq 0\}$, and first two limiting moments denoted by $\overline{X}$, $\overline{X^2}$. It is readily verified that both Markov chain $\{\underline{X}_n, n \geq 0\}$ and $\{X_n, n \geq 0\}$ are positive-recurrent if and only if $\lambda < 1$.

The evaluation of the delay-throughput function under the IRAR II scheme, now follows in the same manner as that for the IRAR I scheme, using Eqs. (3.21)-(3.25) in Eqs. (3.7). The result is summarized by the following Theorem. It is again convenient to represent (3.21) - (3.22) in terms of a single recursive state equation given as,

$$X_{n+1} = N_{n+1} + I_{n+1} I(N_{n+1} \geq 2)$$
$$= I(X_n \geq 2)\left[Z_n(X_n+1) + I_{n+1}\right] + I(X_n \leq 1) Z_{n+1}(1), \qquad (3.26)$$

where $\{Z_n(i)\}$ is defined as before and is statistically independent of $\{I_n\}$. Note the minor difference between (3.26) and (3.15). Probabilities u_o and u_1, defined by (3.10) and satisfying (3.11), can again be evaluated by a repetitive recursive use of the recursive relationship in Fig. 3.4, using (3.14).

Theorem 3.2. For a multi-access channel with R = 0, under an IRAR II scheme, for a packet arrival intensity of $\lambda < 1$, the limiting average packet delay D is given by

$$D = W + 1, \tag{3.27}$$

where

$$\overline{W} = \frac{1}{2}(3\lambda+1) + \frac{1}{2}(1+\lambda)\,\overline{X^2}(\overline{X})^{-1} + (\overline{X})^{-1}[\lambda-2(\lambda+1)u_1] \;. \tag{3.28}$$

The limiting means $\overline{X} = \sum_i iu_i$, $\overline{X^2} = \sum_i i^2u_i$, are given in terms of limiting probabilities u_o, u_1, as follows.

$$\overline{X} = \lambda(1-\lambda)^{-1}(2-u_o-2u_1), \tag{3.29a}$$

$$\overline{X^2} = \lambda(1-\lambda^2)^{-1}[2(1+2\lambda) + (1+4\lambda)\overline{X} - (u_o+u_1)(1+\lambda)-u_1(3+5\lambda)]. \tag{3.29b}$$

For $\lambda \geq 1$, D is unbounded.

Proof. Expressions (3.29a) and (3.29b) are obtained by taking expectations on both sides of Eq. (3.26) and the square of each side of (3.26), respectively, and invoking equilibrium conditions, as in Lemma 3.1. Eq. (3.28) is obtained by calculating $E[W(\underline{X}_n,\underline{X}_{n+1})]$ and using (3.7) and $E[N(\underline{X}_n,X_{n+1})] = \overline{X}$ to obtain

$$\overline{W} = (\overline{X})^{-1}E[W(\underline{X}_n,\underline{X}_{n+1})].$$

In performing the above calculations we note the following relationships.

$$E\{I(X_n \geq 2)N_{n+1}X_n\} = E\{\lambda(X_n+1)X_nI(X_n \geq 2)\} = \lambda(\overline{X^2}-u_1) + \lambda(\overline{X}-u_1) \;;$$

$$E\{N_{n+1}I(N_{n+1} \geq 2)\} = \sum_{i=2}^{\infty} \lambda(i+1)u_i + \lambda(u_o+u_1) - u_1$$

$$= \lambda(\overline{X}-u_1) + \lambda(1-u_o-u_1) + \lambda(u_o+u_1) - u_1$$

$$= \lambda\overline{X} - \lambda u_1 + \lambda - u_1 .$$

Subsequently, we obtain

$$E[W(\underline{X}_n, \underline{X}_{n+1})] = \frac{1}{2}\lambda(\overline{X^2}-u_1) + \frac{1}{2}\lambda(\overline{X}-u_1)$$

$$+ \lambda\overline{X} - \lambda u_1 + \lambda - u_1 + \frac{1}{2}(\overline{X} - u_1) + \frac{1}{2}(\overline{X^2}-u_1)$$

$$= \frac{1}{2}(1+\lambda)\overline{X^2} + \frac{1}{2}(3\lambda + 1)\overline{X} - 2u_1(\lambda+1) + \lambda ,$$

which yields, upon division by $\overline{X}$, expression (3.28). Taking expectations on both sides of (3.26) we obtain, at equilibrium,

$$E(X_{n+1}) = \overline{X} = E(N_{n+1}) + \lambda(1-u_o-u_1)$$

$$= (u_o+u_1)\lambda + \sum_{i=2}^{\infty} u_i\lambda(i+1) + \lambda(1-u_o-u_1)$$

$$= \lambda + \lambda\overline{X} - \lambda u_1,$$

yielding Eq. (3.29a). Taking expectations of the square of each side of (3.26), we have

$$E(X_{n+1}^2) = \overline{X^2} = E(N_{n+1}^2) + E\{I(N_{n+1} \geq 2)I_{n+1}^2\}$$

$$+ 2E\{N_{n+1}I(N_{n+1} \geq 2)I_{n+1}\} ;$$

where

$$E(N_{n+1}^2) = (u_o+u_1)(\lambda^2+\lambda) + \sum_2^{\infty} u_i[\lambda^2(i+1)^2 + \lambda(i+1)]$$

$$= (u_o+u_1)(\lambda+\lambda^2) + \lambda^2(\overline{X^2}-u_1) + (\lambda+2\lambda^2)(\overline{X}-u_1)$$

$$+ (\lambda+\lambda^2)(1-u_o-u_1) ,$$

$$E\{I(N_{n+1} \geq 2)I_{n+1}^2\} = (1-u_o-u_1)(\lambda+\lambda^2) ,$$

$$E\{N_{n+1}I(N_{n+1} \geq 2)I_{n+1}\} = E\{N_{n+1}I(N_{n+1} \geq 2)\} E(I_{n+1})$$

$$= (\lambda\bar{X}-\lambda u_1 + \lambda - u_1)\lambda,$$

yielding subsequently relation (3.29b).

Q.E.D.

The delay-throughput curve for a channel with R = 0, under IRAR II scheme, computed according to (3.27), is shown in Fig. 3.7. We note that the delay-throughput curves under IRAR I and II schemes are essentially the same for $0 \leq \lambda \leq 0.4$, while for $\lambda > 0.4$ IRAR II scheme exhibits a somewhat better performance.

<u>The IRAR III Scheme.</u>

The IRAR III scheme differs from the IRAR II scheme in that it declares the slot following a service period as a reservation slot. We thus expect, when compared to IRAR II scheme, the IRAR III scheme to yield a lower delay-throughput curve at medium and high values of λ, while exhibiting a similar performance at lower values of λ.

A sample function of the channel state process under IRAR III scheme, with R = 0, is shown in Fig. 3.5. Time periods are now defined as follows. A time period constitutes of a single slot if no packets transmit within this slot, or a single (successful) RA transmission has occurred in this slot. Thus, a single slot period will contain a reservation slot with no reservations made in it (to be called a zero reservation slot), or be composed of a RA slot during which no collissions are experienced. A time period which contains more than one slot, will start with either a collission slot or a nonzero reservation slot. In the first case, the period will contain the collission slot and the successive reservation slot and service period. In the second case, the time period will be composed of the nonzero reservation slot and the following service slots.

Considering the n-th time period, we denote the number of transmissions within its first slot again by N_n. If the n-th period contains a collission slot, we let (as for IRAR II scheme) I_n denote the number of new arrivals within this collission slot. We now set X_n to denote the number of reserved transmissions within the n-th period. Thus, for a single slot period we have $X_n = 0$, while $N_n = 0, 1$. Note that the present definition of X_n differs from the corresponding one used for IRAR I, II schemes.

The channel Markov state sequence $\{\underline{X}_n, n \geq 0\}$ is characterized by state $\underline{X}_n$, where $\underline{X}_n = (N_n, I_n, X_{n-1})$. In particular, if $X_{n-1} = 0$ the first slot in the n-th period is a RA slot. On the other hand, if $X_{n-1} > 0$ the first slot of the n-th period is a reservation slot.

The state equations for IRAR III, R = 0, follow the flow diagram of Fig. 3.6, and are presented as follows.

$$X_{n+1} = N_{n+1} + I(X_n=0)\,[I(N_{n+1} \geq 2)I_{n+1} - I(N_{n+1} = 1)] , \quad (3.30a)$$

where

$$N_{n+1} = Z_{n+1}(X_n+1) , \quad (3.30b)$$

and, as before, $\{Z_n(i)\}$ and $\{I_n\}$ are statistically independent sequences of i.i.d. random-variables governed by Poisson distributions with means $i\lambda$ and λ, respectively.

We note from Eqs. (3.30) that $\{X_n, n \geq 0\}$ itself is a Markov chain with stationary transition probabilities. It is again readily shown that both Markov chains are positive-recurrent if and only if $\lambda < 1$. The stationary distribution of $\{X_n\}$ is again denoted by $\{u_i, i \geq 0\}$, and the associated first two moments as $\overline{X} = \sum_i i u_i$ $\overline{X^2} = \sum_i i^2 u_i$. We will note that only the limiting probability of state 0, u_o, is required to evaluate D under IRAR III, when R = 0. Probability u_o can be computed, as for IRAR I, II, by successive iteration governed by recursive relationship (3.30).

To evaluate the packet average waiting-time $\overline{W}$, we need evaluate functions N(•) and W(•). The overall number of packets served during the (n+1)-st period is now given by

$$N(\underline{X}_n, \underline{X}_{n+1}) = X_{n+1} + I(X_n=0)I(N_{n+1}=1). \quad (3.31)$$

The first term on the RHS of (3.31) accounts for the packets served by reservation, while the second term incorporates a single (successful) RA transmission.

The overall sum of waiting-times for packets served during the (n+1)-st period is now given by

$$W(\underline{X}_n,\underline{X}_{n+1}) = \frac{1}{2} X_n X_{n+1} + X_{n+1} + I(X_n=0)X_{n+1} + \frac{1}{2} X_{n+1}(X_{n+1}-1) . \quad (3.32)$$

The first term on the RHS of Eq. (3.32) represents the (averaged, given $\underline{X}_n,\underline{X}_{n+1}$) overall delay of the X_{n+1} packets served within the (n+1)-st period, while waiting through the n-th period, prior to their reservation in the (n+1)-st period. The second and third terms account for the delay of the X_{n+1} packets along their reservation and collission slots, respectively. The fourth term in (3.32) represents the overall (averaged) waiting-time of the X_{n+1} packets after they have made reservations and prior to their admission into service.

The evaluation of the delay-throughput function under the IRAR III scheme now follows in the same manner as those for the IRAR I-II schemes, using Eqs. (3.30) - (3.32) in Eqs. (3.7). The result is summarized by the following theorem.

Theorem 3.3. For a multi-access channel with R = 0, under an IRAR III scheme, for a packet arrival intensity of $\lambda < 1$, the limiting average packet delay D is given by

$$D = W + 1, \quad (3.33)$$

where

$$\overline{W} = [\overline{X}+u_o\lambda e^{-\lambda}]^{-1} \left\{\frac{1}{2} (\lambda+1)\overline{X} + \frac{1}{2} (\lambda+1)\overline{X^2} + u_o\lambda[2-2 \exp(-\lambda) - \lambda \exp(-\lambda)]\right\}, \quad (3.34)$$

and $u_o = \lim_{n \to \infty} P(X_n = 0)$. The limiting means are given by

$$\overline{X} = \lambda(1-\lambda)^{-1} \cdot \left\{1+u_o[1-2\exp(-\lambda)-\lambda\exp(-\lambda)]\right\}, \tag{3.35a}$$

$$\begin{aligned}\overline{X^2} = \lambda(1-\lambda^2)^{-1} \cdot \Big\{&(1+2\lambda)\overline{X} + \lambda+1 - u_o\exp(-\lambda) \\ &+ u_o(1+\lambda)[1-\exp(-\lambda)-\lambda\exp(-\lambda)] \\ &+ 2\lambda u_o[1-\exp(-\lambda)]\Big\}.\end{aligned} \tag{3.35b}$$

For $\lambda \geq 1$, D is unbounded.

<u>Proof</u>. Taking expectations on both sides of (3.30) we obtain, at equilibrium, noting that $E\{N_{n+1}\} = E\{Z(X_n+1)\} = \lambda(\overline{X}+1)$, $E(X_{n+1}) = \overline{X} = \lambda(\overline{X}+1) - u_o\lambda e^{-\lambda} + u_o\lambda(1-e^{-\lambda}-\lambda e^{-\lambda})$, which yields Eq. (3.35a). Taking the square of each side of (3.30), and subsequently the corresponding expections, and noting that

$$E\{N_{n+1}^2\} = E\{Z^2(X_n+1)\} = E\{\lambda^2(X_n+1)^2 + \lambda(X_n+1)\},$$

we obtain, at equilibrium,

$$\begin{aligned}E(X_{n+1}^2) = \overline{X^2} = {}& E\{N_{n+1}^2\} + E\{I(X_n=0)I(N_{n+1}=1)\} \\ &+ E\{I(X_n=0)I(N_{n+1} \geq 2)I_{n+1}^2\} \\ &- 2E\{N_{n+1}I(N_{n+1}=1)I(X_n=0)\} \\ &+ 2E\{N_{n+1}I(X_n=0)I(N_{n+1} \geq 2)I_{n+1}\} \\ = {}& \lambda^2E[(X_n+1)^2] + \lambda E[X_n+1] + E[I(X_n=0)]E[I(N_{n+1}=1)|X_n=0] \\ &+ E(I_{n+1}^2)E[I(X_n=0)]E[I(N_{n+1} \geq 2)|X_n=0] \\ &-2E[I(X_n=0)]E[I(N_{n+1}=1)|X_n=0] \\ &+2E(I_{n+1})E[I(X_n=0)]E[N_{n+1}I(N_{n+1} \geq 2)|X_n = 0] \\ = {}& \lambda^2(\overline{X^2}+2\overline{X}+1) + \lambda(\overline{X}+1) + u_o\lambda\exp(-\lambda) \\ &+ (\lambda^2+\lambda)u_o[1-\exp(-\lambda)-\lambda\exp(-\lambda)] \\ &- 2u_o\lambda\exp(-\lambda) + 2\lambda u_o[\lambda-\lambda\exp(-\lambda)],\end{aligned}$$

which yields Eq. (3.35b). To obtain $\overline{W}$, we first obtain, using (3.31), at equilibrium,

$$E\{N(\underline{X}_n,\underline{X}_{n+1})\} = E(X_{n+1}) + E\{I(X_n=0)I(N_{n+1}=1)\} = \overline{X} + u_o\lambda \exp(-\lambda).$$

Using (3.32), we have at equilibrium,

$$E\{W(\underline{X}_n,\underline{X}_{n+1})\} = \frac{1}{2} E\{X_n X_{n+1}\} + E(X_{n+1}) + E\{I(X_n=0)X_{n+1}\}$$

$$+ \frac{1}{2} E\{X_{n+1}(X_{n+1}-1)\}$$

$$= \frac{1}{2} (\lambda\overline{X}+\lambda\overline{X^2}) + \overline{X} + u_o\lambda[2-2 \exp(-\lambda)-\lambda \exp(-\lambda)]$$

$$+ \frac{1}{2} [\overline{X^2}-\overline{X}] = \frac{1}{2} \overline{X}(\lambda+1) + \frac{1}{2} \overline{X^2}(\lambda+1) + u_o\lambda[2-2 \exp(-\lambda)-\lambda \exp(-\lambda)],$$

noting that

$$E[X_n X_{n+1}] = E\{X_n E[X_{n+1}|X_n]\} = E\{X_n\lambda(X_n+1)\} = \lambda(\overline{X^2}+\overline{X}) ,$$

and that by (3.30),

$$E\{I(X_n=0)X_{n+1}\} = E\{I(X_n=0)N_{n+1}\} - E\{I(X_n=0)I(N_{n+1}=1)\}$$

$$+ E\{I(X_n=0)I(N_{n+1} \geq 2)I_{n+1}\}$$

$$= u_o\lambda - u_o\lambda \exp(-\lambda) + u_o\lambda[1-\exp(-\lambda)-\lambda \exp(-\lambda)] .$$

Dividing now $E\{W(\underline{X}_n,\underline{X}_{n+1})\}$ by $E\{N(\underline{X}_n,\underline{X}_{n+1})\}$, we obtain Eq. (3.34). The recurrence relationships for the moments $\overline{X}$, $\overline{X^2}$, readily show that D is unbounded for $\lambda \geq 1$.

Q.E.D.

Delay Throughput Performance Comparisons

The delay-throughput performance curves for a multi-access communication channel with R = 0, under IRAR I, II, III schemes, are shown in Fig. 3.7. We note that the IRAR III scheme exhibits a better performance curve, over the whole throughput range, when compared to IRAR I-II performance curves. In particular, IRAR III yields a significant improvement in performance within the higher throughput range, $\lambda > 0.5$, as expected. For $\lambda < 0.5$, the performance curves obtained by IRAR I, II, III schemes are very close.

We also show in Fig. 3.7 the delay-throughput curve for the ARDA I reservation scheme. Over the throughput range $0 \le \lambda \le 0.4$ we note IRAR I, II, III schemes to yield lower packet delay values than those obtained by an ARDA I scheme. This is due to the extra reservation delay required by the ARDA I scheme, when compared with the incorporation of the RA operation in the IRAR scheme. For higher throughput values, $\lambda > 0.4$, we find ARDA I scheme to yield a delay-throughput curve which is below those obtained by IRAR I-II schemes, and somewhat above the IRAR III performance curve.

We thus conclude that the IRAR III scheme yields an excellent delay-throughput performance curve, uniformly (over the whole throughput range) better than the corresponding curves obtained by ARDA I or IRAR I-II schemes. Performance characteristics close to those obtained by IRAR III scheme are also exhibited by IRAR I-II scheme for $\lambda < 0.5$, and by ARDA I scheme for $\lambda > 0.5$.

In Fig. 3.8 we show curves representing probabilities u_o and u_1 vs. λ, under IRAR I, II, III schemes, when R = 0. Note that, by definition of the corresponding state, probabilities u_o and u_1 represent the probabilities of no transmission and a single (successful RA) transmission at a slot for IRAR I-II schemes. For an IRAR III scheme, however, u_o and u_1 represent the probabilities that no reservations and a single reservation, respectively, will be made at any period. We further note that if we incorporate an (analytical or empirical) estimate of u_o, u_1, the delay-throughput equations for an IRAR scheme are fully analytically given by the expressions presented in this section.

In Fig. 3.9 we show curves representing the RA-to-reservation ratio vs. λ, under IRAR I-II schemes. The latter ratio is defined as the ratio between the average numbers of packets transmitted without reservation and those transmitted following a reservation operation. Curves are shown for R = 0, 1, 12. We note that for $0 \le \lambda \le 0.4$, this ratio decreases rather linearly with λ, while for

$\lambda > 0.4$ an almost exponential decrease with λ is noted for $R = 0$. For higher values of R, R=1, 12, the decrease of this ratio with $\lambda > 0.4$, is much slower. This is explained by noting that at higher channel propagation delay values, more slots are available for RA transmissions between a reservation slot and the successive corresponding service period, as shown in the next section.

IV. Performance Analysis of IRAR Schemes for Channels with $R \geq 1$

The IRAR IV-V Schemes

We consider now multi-access communication channels with longer propagation delay values, $R \geq 1$. As for the $R = 0$ channel, the delay-throughput performance curves under the IRAR III scheme are noted to be uniformly (over the whole throughput range) better than those obtained under IRAR I-II schemes. It is thus of main interest here to derive the delay-throughput characteristics under an IRAR III scheme. In particular, we are interested in the latter characteristics for medium to large network throughput values, since for low throughput values all IRAR schemes yield performance curves very close to those obtained under a pure random-access scheme.

The analysis of an IRAR III (or I-II) scheme for $R \geq 1$, is performed by following the same procedure presented in Section III for $R = 0$ channels. However, the underlying channel state process will now include additional parameters (to principally indicate whether a group starts with a RA slot or a reservation slot, the latter to be used by packets colliding in the previous slot). The Markov state sequence will thus assume now a more complexed structure, although the technique for evaluating D remains the same.

To provide a simpler procedure for evaluating the performance characteristics of an IRAR III scheme for $R \geq 1$, we present here the IRAR IV scheme. This scheme, which is of interest due to its own merits, will yield delay-throughput curves which are very close to those associated with the IRAR III scheme over the medium to large network throughput values, and are only somewhat above the latter curves for lower throughput values. These resulting characteristics are readily observed while comparing the protocols of these schemes and reviewing typical sample functions of the channel processes under these schemes. In stating the protocol of the IRAR IV scheme, an idle slot is defined as a slot during which no transmissions are allowed. A single-slot period is defined, as in Section III, to include a zero reservation slot (containing no reservations), or a single non-collission RA slot, prior to which the network contains no waiting unserved packets.

Protocol for IRAR IV Scheme

1. A collission slot which follows a single-slot period, is followed by (R-1) successive RA slots and R successive idle slots. Otherwise, any slot which has not been reserved as a service or reservation slot is declared to be a random-access (RA) slot.

2. The (first unreserved) slot following the instant at which a collission is recognized by network terminals, is established as a reservation slot. Also, any (unreserved) slot following a service slot is declared as a reservation slot.

3. Service slots are established (at unreserved slots) for reserved messages immediately following the reception of the corresponding reservation packet, in accordance with the underlying service ordering discipline.

4. Upon its arrival, a newly arriving message checks whether there exists any reservation slot preceding the first available RA slot. If the latter is the case, the message sends a reservation packet in the latter reservation slot. Otherwise, the information packet is transmitted in the first available RA slot. Then, in the latter situation, if collision is subsequently recognized, reservation is made for this packet in the first available reservation slot.

5. A service period is set to last for at least R slots.

We note that IRAR IV scheme is governed by a protocol which is identical to that of an IRAR III scheme, except for the first part of instruction 1 and the introduction of instruction 5 in the IRAR V protocol. The first part of instruction 1 sets up a sequence of consecutive (R-1) RA slots and R idle slots to follow a collission slot which follows a single-slot period. Subsequently, all colliding packets trying RA transmissions within the above declared R RA slots, will be able to make reservations within the reservation slot which is set-up to follow the last idle slot (see Fig. 4.1). This procedure much reduces the complexity of the state process following a non-collission RA slot. It is further observed to have little effect on the delay-throughput curve, in particular at medium and large throughput values.

Instruction 5 sets up each service period to be no shorter than R slots. Idle slots need to be added to the service period, if there are less than R service slots. This allows for all the colliding packets transmitting previous to the beginning of the service period to make reservations at a reservation slot declared at the slot following immediately the last service slot within this service period. This requirement yields, as well, a simplified structured for the channel state process. However, it results with a somewhat higher delay-throughput curve (due to the insertion of the service idle slots), when $R \geq 2$. (For $R = 1$, instruction 5 is clearly not needed.)

To analytically reduce the effect of instruction 5 on the resulting delay-throughput curve, and thus obtain a closer approximation to the corresponding curve obtained by an IRAR III scheme, we can make the following assumption. We assume each service period to contain only its actual service slots (not adding any idle slots to it), but require the reservation slot, which follows this service period, to contain the reservations made by any of the packets colliding during the R slots prior to this period (see Fig. 4.1). The resulting scheme is then called IRAR IV_A. We note that the latter scheme is not, in general, realizable, since a propagation delay of less than R slots is sometimes assumed (to allow for the propagation of collission information over less than R slots, prior to the reservation slot, when a service period shorter than R slots occurs). However, scheme IV_A will yield delay-throughput curves which are very close to those obtained under an IRAR III scheme (since the latter two schemes exhibit very similar sample function behavior for the underlying state process, mainly under medium and high network throughput values).

A simple extension of the IRAR IV scheme, associated with a modification of the first part of instruction 1, yields IRAR V scheme.

Protocol for IRAR V Scheme: Instructions 2-5 of IRAR V are identical to the corresponding instructions 2-5 for IRAR IV. Instruction 1 for IRAR V is given as follows: 1) Following every single-slot period we declare the next successive R slots as RA slots, the next following successive R slots as idle slots and the next following slot as a reservation slot. Otherwise, any slot which has not been reserved as a service or reservation slot is declared to be a RA slot. |

Thus, IRAR V protocol is identical to the IRAR IV protocol, except that the initial period, following a single-slot period and containing R RA slots, R idle slots and a reservation slot, starts with a collission slot for an IRAR IV scheme, while starting with any (RA) slot for an IRAR V scheme. (Therefore, under an IRAR V scheme, single-slot periods constitute always of a single reservation slot containing no reservations.) An IRAR V_A scheme is defined in a manner analogous to that used to define an IRAR IV_A procedure.

We note that scheme IRAR V will yield a delay-throughput performance curve somewhat higher than that obtained under an IRAR IV scheme, due to more fixed (non-adaptive) scheduling associated with the setting of the initial period, following a single-slot period. We, however, present here the performance of the channel also under an IRAR V scheme, since the latter (more-fixed) structure can be desirable in various actual situations, while yielding a performance curve quite close to that obtained under an IRAR IV scheme.

Performance Analysis

We consider now a channel with $R \geq 1$ under an IRAR IV scheme. The channel state stochastic sequence $\{\underline{X}_n\}$ is characterized as follows. We note the channel process to contain three types of time periods (see Fig. 4.1 for an example of a sample function). The first type of time period is the single-slot period defined above, containing a single zero reservation slot or a single non-collission RA slot, at which time the network contains no waiting unserved packets. The second type of a time period is the period following a single-slot period, called an initial period. (Note that a slot in a single-slot period constitutes a regeneration point for the channel state stochastic process.) By instruction 1 of the protocol for IRAR IV scheme, the initial period is set to contain successively, as its first 2R+1 slots, R RA slots, R idle slots and a reservation slot; the first slot being always a collission slot. Included in the initial group are also R RA slots which follow the reservation slot (during which reservation information is broadcasted) and the service period following the latter RA slots (see Fig. 4.1). The third type of period, called regular period, consists of a nonzero reservation slot as a first slot, assuming this reservation slot does not follow a period of idle slots (in which case we have an initial period), followed by R RA slots (during which reservation information is broadcasted) and a service period. In Fig. 4.1, we observe a sample function representing the successive occurrence of a single-slot period, an initial period and a regular period. For medium and high network throughput values, the successive occurrence of regular periods is the most dominant feature of a sample function of the channel state process. (For these throughput values, a similar observation can be made for a channel under an IRAR III scheme.)

To analytically represent the channel state sequence $\{\underline{X}_n\}$, we make the following state definitions, for the n-th time period. We set N_n and I_n to denote the number of packets making reservations, within the n-th period, after experiencing collissions and the number of newly arriving packets making reservations, respectively. We then let X_n denote the number of reserved transmissions within the n-th group. The number of transmissions within the first RA slot of the n-th group is denoted by $N_n^{(1)}$ (if such a slot exists; otherwise, we set $N_n^{(1)} = 0$).

Thus, if the n-th period is a single-slot period we have $N_n = I_n = X_n = 0$. If the latter period consists of an RA slot (so that $X_{n-1} = 0$), then $N_n^{(1)} = 0$ or 1. Otherwise, this is a zero reservation slot, and $X_{n-1} > 0$. If the n-th period is an initial or regular period, we have

$$X_n = N_n + I_n .$$

The variable N_n then represents the overall number of collissions occurring within the previous group of R RA slots. Variables I_n represents the overall number of new arrivals with the previous service period, or during the previous R idle slots, when considering a regular or initial period, respectively. Thus, the state of the Markov channel state process is set up as

$$\underline{X}_n = (N_n, I_n, N_n^{(1)}, X_{n-1}) . \tag{4.1}$$

The flow diagram describing the transition probability function of the Markov state sequence, is shown in Fig. 4.2. The state equations are thus readily seen to be represented as follows.

$$X_{n+1} = [I(X_n > 0) + I(X_n = 0) I(N_{n+1}^{(1)} \geq 2)] [N_{n+1} + I_{n+1}] , \tag{4.2}$$

where

$$N_{n+1} = \sum_{i=1}^{R} N_{n+1}^{(i)} I(N_{n+1}^{(i)} \geq 2) , \tag{4.3}$$

$$I_{n+1} \equiv \begin{cases} Z_{n+1}(X_n+1), & \text{if } X_n \geq \delta \\ Z_{n+1}(R+1), & \text{if } X_n < \delta , \end{cases} \tag{4.4}$$

and $\{N_n^{(i)}\}$, $\{Z_n(j)\}$ are independent i.i.d. sequences governed by a Poisson distribution with mean λ and $j\lambda$, respectively, for each $j \geq 0$. Furthermore, we set

$$\delta = \begin{cases} 1, & \text{for IRAR IV}_A \\ R, & \text{for IRAR IV.} \end{cases} \tag{4.5}$$

State equations (4.2)-(4.5) are explained as follows. For $\{X_n=0, N_{n+1}^{(1)} \leq 1\}$, the (n+1)-st period consists of a single non-collission RA slot, and $X_{n+1} = 0$, as indicated by (4.2). In all other cases, we have $X_{n+1} = N_{n+1} + I_{n+1}$. Variable N_{n+1} is given by (4.3), where $N_{n+1}^{(i)}$ is set to indicate the number of transmissions within the i-th slot in the corresponding group of R RA slots. Variable I_{n+1} represents the overall number of (Poisson distributed) new arrivals during the n-th service period, when a regular (n+1)-st period is considered. For an initial period, I_{n+1} represents the overall number of new arrivals during the group of R idle slots. Thus, for an IRAR IV scheme we obtain relation (4.4) with $\delta = R$. For an IRAR IV_A scheme we set $\delta = 1$ in (4.4), so that we have

$$I_{n+1} = I(X_n=0)Z_{n+1}(R+1) + I(X_n>0)Z_{n+1}(X_n+1) . \tag{4.6}$$

Thus, for IRAR IV_A we assume the number I_{n+1} of new packets reserving in the (n+1)-st period to arrive during the previous X_n service slots if $X_n > 0$, or during the previous R idle slots if $X_n = 0$ (and the (n+1)-st period is an initial period).

We note that the stochastic sequence $\{X_n, n \geq 0\}$, where X_n denotes the number of reserved transmissions within the n-th slot, is a Markov sequence with homogeneous transition probabilities characterized by Eqs. (4.2)-(4.6). It is readily noted that $\{X_n\}$ is positive-recurrent if and only if $\lambda < 1$. For $\lambda < 1$, we set $\{u_i, i \geq 0\}$ to denote the steady-state distribution,

$$u_i = \lim_{n \to \infty} P(X_n = i),\ i \geq 0\ , \tag{4.7}$$

and $\bar{X}$, $\overline{X^2}$ as the limiting moments,

$$\bar{X} = \sum_i i u_i\ ,\quad \overline{X^2} = \sum i^2 u_i\ . \tag{4.8}$$

For explicit analytical calculations, we will consider henceforth an IRAR IV_A scheme. The latter will be shown to yield a performance curve which is closer to that obtained under an IRAR III scheme, than the one resulting under an IRAR IV scheme. We will however indicate also the procedure for evaluating the performance curve under an IRAR IV scheme.

Using recurrence relationships (4.2) - (4.6), we obtain the moments in (4.8) in terms of the limiting probability of state 0, u_o.

<u>Lemma 4.1</u>. For a multi-access channel with $R \geq 1$, under an IRAR IV_A scheme, for $\lambda < 1$, the limiting means $\bar{X}$, $\overline{X^2}$ are given by

$$\bar{X} = (1-\lambda)^{-1}\ \{(1-u_o)\lambda[1+R(1-\lambda)] + u_o\lambda(R+1)[1-\exp(-\lambda)-\lambda\ \exp(-\lambda)] +$$

$$u_o\lambda[1-\exp(-\lambda)][1+(R-1)(1-\exp(-\lambda)-\lambda\ \exp(-\lambda))]\}\ ; \tag{4.9}$$

$$\overline{X^2} = (1-\lambda^2)^{-1}(A_3+u_oA_2)\ , \tag{4.10}$$

where

$$A_3 = (1-u_o)R[\lambda^2+\lambda-\lambda\exp(-\lambda)] + (1-u_o)R(R-1)\lambda^2[1-\exp(-\lambda)]^2$$
$$+ 2R\lambda^2[1-\exp(-\lambda)](1-u_o) + (\lambda^2+\lambda)(1-u_o)$$
$$+ \bar{X}\{2R\lambda^2[1-\exp(-\lambda)] + \lambda + 2\lambda^2\} , \qquad (4.11a)$$

$$A_2 = A_1 + 2\lambda[1-\exp(-\lambda)]\{1+(R-1)[1-\exp(-\lambda)-\lambda\exp(-\lambda)]\}$$
$$+ [1-\exp(-\lambda)-\lambda \exp(-\lambda)][\lambda^2(R+1)^2 + \lambda(R+1)], \qquad (4.11b)$$

$$A_1 = \lambda^2 + \lambda-\lambda \exp(-\lambda) + 2\lambda[1-\exp(-\lambda)](R-1)\lambda[1-\exp(-\lambda)]$$
$$+ (R-1)[\lambda^2 + \lambda - \lambda \exp(-\lambda)]$$
$$+ (R-1)(R-2)\lambda^2[1-\exp(-\lambda)]^2 , \qquad (4.11c)$$

where $u_o = \lim_{n \to \infty} P(X_n=0)$.

<u>Proof</u>. Taking expectation on both sides of (4.2), and using (4.3), (4.6), we obtain at equilibrium,

$$E(X_{n+1}) = \bar{X} = E\{[I(X_n>0)] \cdot [N_{n+1} + Z_{n+1}(X_n+1)]\}$$
$$+ E[I(X_n=0)]E\{I(N_{n+1}^{(1)} \geq 2)[N_{n+1} + Z_{n+1}(R+1)]$$
$$= (1-u_o)E(N_{n+1}) + \lambda E\{I(X_n > 0)(X_n + 1)\}$$
$$+ u_o E\{N_{n+1} I(N_{n+1}^{(1)} \geq 2)\} + u_o\lambda(R+1)[1-\exp(-\lambda)-\lambda\exp(-\lambda)]$$
$$= (1-u_o)R\lambda[1-\exp(-\lambda)] + \lambda(1-u_o) + \lambda\bar{X}$$
$$+ u_o\{\lambda[1-\exp(-\lambda)] + (R-1)\lambda[1-\exp(-\lambda)][1-\exp(-\lambda)$$
$$- \lambda \exp(-\lambda)]\} + u_o\lambda(R+1)[1-\exp(-\lambda)-\lambda \exp(-\lambda)] , \qquad (4.12)$$

since

$$E(N_{n+1}) = R[\lambda-\lambda\exp(-\lambda)] , \qquad (4.13)$$

$$E\{N_{n+1} I(N_{n+1}^{(1)} \geq 2)\} = E\{N_{n+1}^{(1)} I(N_{n+1}^{(1)} \geq 2)\}$$

$$+ E\{I(N_{n+1}^{(1)} \geq 2)\} \; E\{\sum_{i=2}^{R} N_{n+1}^{(i)} I(N_{n+1}^{(i)} \geq 2)\}$$

$$= \lambda[1-\exp(-\lambda)] + [1-\exp(-\lambda)-\lambda \exp(-\lambda)](R-1)\lambda[1-\exp(-\lambda)] \; . \quad (4.14)$$

Solving Eq. (4.12) for $\bar{X}$ we obtain Eq. (4.9).

To obtain Eq. (4.10) we square both sides of (4.2) and take expectations, assuming equilibrium. We obtain then,

$$E(X_{n+1}^2) = \overline{X^2} = E\{I(X_n > 0)[N_{n+1} + Z_{n+1}(X_n+1)]^2\}$$

$$+ E\{I(X_n=0) I(N_{n+1}^{(1)} \geq 2)[N_{n+1} + Z_{n+1}(R+1)]^2\} \; . \quad (4.15)$$

The first expression is given by

$$E\{I(X_n > 0)[N_{n+1} + Z_{n+1}(X_n+1)]^2\}$$

$$= (1-u_o)E(N_{n+1}^2) + 2E(N_{n+1})\lambda E\{I(X_n > 0)(X_n+1)\}$$

$$+ E\{I(X_n > 0)[\lambda^2(X_n+1)^2 + \lambda(X_n+1)]\}$$

$$= (1-u_o)R[\lambda^2 + \lambda - \lambda \exp(-\lambda)] + (1-u_o)(R^2-R)\lambda^2[1-\exp(-\lambda)]^2$$

$$+ 2R\lambda[1-\exp(-\lambda)][\lambda\bar{X} + \lambda(1-u_o)]$$

$$+ (\lambda^2+\lambda)(1-u_o) + (\lambda+2\lambda^2)\bar{X} + \lambda^2\overline{X^2} \triangleq \lambda^2\overline{X^2} + A_3 \; , \quad (4.16)$$

where A_3 is given by (4.11a). To obtain (4.16), we note that

$$E(N_{n+1}^2) = E\{[\sum_{i=1}^{R} N_{n+1}^{(i)} I(N_{n+1}^{(i)} \geq 2)]^2\}$$

$$= RE\{I(N_{n+1}^{(1)} \geq 2)[N_{n+1}^{(1)}]^2\}$$

$$+ (R^2-R)\{E[N_{n+1}^{(1)} I(N_{n+1}^{(1)} \geq 2)]\}^2$$

$$= R[\lambda^2 + \lambda - \lambda \exp(-\lambda)] + (R^2-R)[\lambda - \lambda \exp(-\lambda)]^2. \quad (4.17)$$

To evaluate the second expression of (4.15), we use (4.14) and note that

$$E\{N_{n+1}^2 I(N_{n+1}^{(1)} \geq 2)\} = E\{I(N_{n+1}^{(1)} \geq 2)[N_{n+1}^{(1)}]^2\}$$

$$+ 2E\{I(N_{n+1}^{(1)} \geq 2)N_{n+1}^{(1)}\} (R-1)E\{N_{n+1}^{(2)} I(N_{n+1}^{(2)} \geq 2)\}$$

$$+ E\{[\sum_{i=2}^{R} N_{n+1}^{(i)} I(N_{n+1}^{(i)} \geq 2)]^2\} = A_1 , \qquad (4.18)$$

where A_1 is given by (4.11c). Finally, using (4.14), (4.18), we obtain the second expression of (4.15) to be given by

$$E\{I(X_n=0)I(N_{n+1}^{(1)} \geq 2)[N_{n+1} + Z_{n+1}(R+1)]^2\}$$

$$= u_o A_1 + 2u_o\{\lambda[1-\exp(-\lambda)][1+(R-1)(1-\exp(-\lambda)-\lambda \exp(-\lambda))]\}$$

$$+ u_o[1-\exp(-\lambda)-\lambda \exp(-\lambda)][\lambda^2(R+1)^2 + \lambda(R+1)] = u_o A_2 , \qquad (4.19)$$

where A_2 is given by (4.11b). Substituting Eqs. (4.16) and (4.19) into Eq. (4.15), we obtain an equation for $\overline{X^2}$, which when solved yields Eq. (4.10).

Q.E.D.

To obtain the packet average waiting-time function $\overline{W}$ under IRAR IV-IV$_A$ schemes, we need first evaluate the associated N(•), W(•) functions. To express N(•), we set

$$J_n = \sum_{i=1}^{R} I(N_{n+1}^{(i)} = 1) , \qquad (4.20)$$

to denote the overall number of RA successful transmissions within the n-th period. (Note that we set $N_n^{(i)} = 0$, if the n-th period contains no i-th RA slot.) The overall number of packets served (successfully transmitted) during the (n+1)-st period is then given by

$$N(\underline{X}_n, \underline{X}_{n+1}) = X_{n+1} + I(X_n=0)I(N_{n+1}^{(1)} = 1)$$

$$+ I(X_n=0)I(X_{n+1} > 0)[J_{n+1} + J_{n+2}] + J_{n+2}I(X_n > 0)I(X_{n+1} > 0). \qquad (4.21)$$

The first term on the RHS of (4.21), X_{n+1}, represents the overall number of packets served by reservation during the (n+1)-st period. The second term accounts for a successful RA transmission within a single-slot (n+1)-st period. The third term represents the overall number of successful RA transmissions within an initial (n+1)-st period. The last terms yields the overall number of successful RA transmissions within a regular (n+1)-st period.

The overall sum of waiting times for packets served during the (n+1)-st period is given as follows.

$$W(\underline{X}_n, \underline{X}_{n+1}) = [I(X_n > 0) + I(X_n = 0) I(N_{n+1}^{(1)} \geq 2)]$$

$$\cdot \left\{ \frac{1}{2} N_{n+1}(R+1) + N_{n+1}(X_n \Lambda R) + \frac{1}{2} I_{n+1}(X_n \Lambda R) \right.$$

$$\left. + X_{n+1}(1+R) + \frac{1}{2} X_{n+1}(X_{n+1}-1) \right\} , \qquad (4.22)$$

where

$$(X_n \Lambda R) \triangleq \text{Max}(X_n, R), \text{ for IRAR IV;} \qquad (4.23a)$$

$$(X_n \Lambda R) \triangleq \begin{cases} R , & \text{if } X_n = 0 \\ X_n , & \text{if } X_n > 0 \end{cases} \quad \text{for IRAR IV}_A . \qquad (4.23b)$$

The indicator-function term in (4.22) sets the waiting-time function equal to zero for a non-collission single RA slot. The next first term, $\frac{1}{2} N_{n+1}(R+1)$, represents the overall delay (averaged w.r.t. packet times of arrival) of the N_{n+1} packets over the R-slot period within which they have collided. The same N_{n+1} packets experience a further delay while waiting for the termination of the previous service period. The latter overall delay is thus given as $N_{n+1}(X_n \Lambda R)$, where function $(X_n \Lambda R)$ is given by (4.23a) for the IRAR IV scheme (so that a minimum service period length of R slots is imposed), and by (4.23b) for the IRAR IV_A scheme. Similarly, the term $\frac{1}{2} I_{n+1}(X_n \Lambda R)$

expresses the overall delay (averaged w.r.t. packet times of arrival) of the I_{n+1} packets during their period of arrival, which is $(X_n \wedge R)$ slots long. (Note Eq. (4.6) when considering the IRAR IV_A scheme.) The term $X_{n+1}(1+R)$ accounts for the overall delay of the X_{n+1} reserving packets over the reservation slot and the R propagation slots. The last term in (4.22), $\frac{1}{2} X_{n+1}(X_{n+1}-1)$, represents the overall sum of waiting-times of the X_{n+1} reserving packets, following their acknowledged broadcasted reservation and prior to their admission into service. (As in Section III, we can assume a random ordering, first-come first-served, fixed priority or any other service discipline, to control the order of service of these X_{n+1} packets, as long as the overall sum of waiting-time remains unchanged.)

For an IRAR IV_A scheme, we now use Eqs. (4.21)-(4.22) and the statistical characteristics of the Markov state process, Eqs. (4.2)-(4.6), to analytically evaluate the average packet waiting-time. The result is presented by the following theorem.

Theorem 4.1. For a multi-access channel under an IRAR IV_A scheme, for $R \geq 1$ and $\lambda < 1$, the limiting average packet delay D is given by

$$D = \bar{W} + 1 + R \,, \tag{4.24}$$

where

$$\bar{W} = \{E[W(\underline{X}_n, \underline{X}_{n+1})]\}\{E[N(\underline{X}_n, \underline{X}_{n+1})]\}^{-1}, \tag{4.25}$$

and

$$\begin{aligned} E[W(\underline{X}_n, \underline{X}_{n+1})] = {} & \bar{X}\{R\lambda[1-\exp(-\lambda)] + \tfrac{1}{2} + R + \tfrac{1}{2}\lambda\} \\ & + \tfrac{1}{2}\overline{X^2}(1+\lambda) + \tfrac{1}{2} R\lambda[1-\exp(-\lambda)](R+1)(1-u_o) \\ & + \lambda[1-\exp(-\lambda)]\{1+(R-1)[1-\exp(-\lambda)-\lambda\exp(-\lambda)]\} \\ & \cdot \left\{\tfrac{3}{2} u_o R + \tfrac{1}{2} u_o\right\} + \tfrac{1}{2}\lambda R(R+1)u_o[1-\exp(-\lambda)-\lambda\exp(-\lambda)] \; ; \end{aligned} \tag{4.26}$$

$$E[N(\underline{X}_n,\underline{X}_{n+1})] = \overline{X} + R(1-u_o)\lambda \exp(-\lambda)$$

$$+ (R-1)\lambda \exp(-\lambda)u_o[1-\exp(-\lambda)-\lambda \exp(-\lambda)] + \lambda u_o \exp(-\lambda). \quad (4.26)$$

The limiting moments $\overline{X}$, $\overline{X^2}$ are given by Eqs. (4.9)-(4.11), and $u_o = \lim_{n\to\infty} P(X_n=0)$. For $\lambda \geq 1$, D is unbounded.

<u>Proof</u>. Eq. (4.25) follows by (3.7), noting again the $\{\underline{X}_n\}$ is positive-recurrent if and only if $\lambda < 1$. To obtain Eq. (4.27) we take the limiting expectation of Eq. (4.21). We obtain,

$$E[N(\underline{X}_n,\underline{X}_{n+1})] = E(X_{n+1}) + E[I(X_n=0)]E[I(N^{(1)}_{n+1} = 1)]$$

$$+ E[I(X_n=0)I(X_{n+1} > 0)][E(J_{n+1}|X_n=0,X_{n+1} > 0) + E(J_{n+2})]$$

$$+ E[I(X_n > 0)I(X_{n+1} > 0)]E(J_{n+2}). \quad (4.28)$$

To compute (4.28) we note that, at equilibrium, we have

$$E(J_{n+1}) = E(J_{n+2}) = E\left\{\sum_{i=1}^{R} I(N^{(i)}_{n+1} = 1)\right\} = R\lambda \exp(-\lambda); \quad (4.29)$$

$$E[I(X_n=0)I(X_{n+1}>0)] = E[I(X_n=0)] - E[I(X_n=0)I(X_{n+1}=0)]$$

$$= u_o-u_oP(X_{n+1}=0|X_n=0) = u_o-u_o[\exp(-\lambda) + \lambda \exp(-\lambda)]$$

$$= u_o[1-\exp(-\lambda) - \lambda \exp(-\lambda)] \quad (4.30)$$

$$E[I(X_n>0)I(X_{n+1}>0)] = E[I(X_{n+1}>0)] - E[I(X_{n+1}>0)I(X_n=0)]$$

$$= (1-u_o) - u_o[1-\exp(-\lambda) - \lambda \exp(-\lambda)]; \quad (4.31)$$

$$E[J_{n+1}|X_n=0, X_{n+1}>0] = E\left[\sum_{i=2}^{R} I(N^{(i)}_{n+1} = 1)\right] = (R-1)\lambda \exp(-\lambda). \quad (4.32)$$

Substituting (4.29)-(4.32) into (4.28) we obtain

$$E[N(\underline{X}_n,\underline{X}_{n+1})] = \overline{X} + u_o\lambda\exp(-\lambda) + u_o[1-\exp(-\lambda)-\lambda\exp(-\lambda)](2R-1)$$

$$\cdot\ \lambda\exp(-\lambda) + R\lambda\exp(-\lambda)\{(1-u_o)-u_o[1-\exp(-\lambda)-\lambda\exp(-\lambda)]\},$$

which yields Eq. (4.27).

To obtain Eq. (4.26) we take the limiting expectation of (4.22). For that purpose, we note that, at equilibrium, the following relationships are obtained.

$$E\{[I(X_n > 0) + I(X_n = 0)I(N^{(1)}_{n+1} \geq 2)]N_{n+1}(X_n \wedge R)\}$$

$$= E\{I(X_n > 0)N_{n+1}X_n + I(X_n = 0)I(N^{(1)}_{n+1} \geq 2)N_{n+1}R\}$$

$$= \overline{X}E(N_{n+1}) + u_oRE\{N_{n+1}I(N^{(1)}_{n+1} \geq 2)\}\ , \tag{4.33}$$

where $E(N_{n+1})$ and $E\{N_{n+1}I(N^{(1)}_{n+1} \geq 2)\}$ are given by (4.13) and (4.14), respectively. Also, using (4.6), we have

$$E\{[I(X_n > 0) + I(X_n = 0)I(N^{(1)}_{n+1} \geq 2)]\ \frac{1}{2}\ I_{n+1}(X_n \wedge R)\}$$

$$= E\{I(X_n > 0)\ \frac{1}{2}\ Z_{n+1}(X_n + 1)X_n\} + E\{I(X_n = 0)I(N^{(1)}_{n+1} \geq 2)\ \frac{1}{2}\ RZ_{n+1}(R+1)\}$$

$$= \frac{1}{2}\ E\{\lambda X_n(X_n + 1)\} + \frac{1}{2}\ \lambda R(R+1)E\{I(X_n = 0)I(N^{(1)}_{n+1} \geq 2)\}$$

$$= \frac{1}{2}\ \lambda\overline{X} + \frac{1}{2}\ \lambda\overline{X^2} + \frac{1}{2}\ \lambda R(R+1)u_o[1-\exp(-\lambda)-\lambda\exp(-\lambda)]. \tag{4.34}$$

Using (4.33)-(4.34) in computing the limiting mean of (4.22), we obtain,

$$E[W(\underline{X}_n,\underline{X}_{n+1})] = \frac{1}{2}\ (R+1)\Big[(1-u_o)E(N_{n+1}) + u_oE[N_{n+1}I(N^{(1)}_{n+1} \geq 2)]\Big]$$

$$+\ \overline{X}E(N_{n+1}) + u_oRE[N^{(1)}_{n+1} \geq 2)] + \frac{1}{2}\ \lambda\overline{X} + \frac{1}{2}\ \lambda\overline{X^2}$$

$$+\ \frac{1}{2}\ \lambda R(R+1)u_o[1-\exp(-\lambda)-\lambda\exp(-\lambda)] + \overline{X}(1+R)$$

$$+\ \frac{1}{2}\ \overline{X^2} - \frac{1}{2}\ \overline{X}\ . \tag{4.35}$$

Substituting (4.13)-(4.14) into (4.35), we derive Eq. (4.26).

Q.E.D.

The analysis of IRAR V-V_A schemes is analogous to that presented above for IRAR IV-IV_A schemes. The underlying Markov state sequence $\{\underline{X}_n\}$ is defined similarly, and is now governed by the transition probability function described in Fig. 4.3. The delay-throughput characteristics are summarized by the following Lemma and Theorem. (Proofs are similar to those presented above and, therefore, are not presented here.)

<u>Lemma 4.1.</u> For a multi-access channel under IRAR V-V_A schemes, with $R \geq 1$, the average packet delay D is given by

$$D = \overline{W} + 1 + R , \tag{4.36}$$

with $\overline{W}$ expressed as the (w.p.1) limit

$$\overline{W} = \lim_{M \to \infty} \frac{\sum_{n=1}^{M} W(\underline{X}_n, \underline{X}_{n+1})}{\sum_{n=1}^{M} N(\underline{X}_n, \underline{X}_{n+1})} , \tag{4.37}$$

and

$$W(\underline{X}_n, \underline{X}_{n+1}) = \frac{1}{2} N_{n+1}(R+1) + N_{n+1}(X_n \wedge R) + \frac{1}{2} I_{n+1}(X_n \wedge R)$$

$$+ X_{n+1}(1+R) + \frac{1}{2} X_{n+1}(X_{n+1}-1) , \tag{4.38}$$

$$N(\underline{X}_n, \underline{X}_{n+1}) = X_{n+1} + J_{n+2} I(X_{n+1} > 0) + J_{n+1} I(X_n = 0), \tag{4.39}$$

where $(X_n \wedge R)$ is given by (4.23). |

Using expressions (4.37)-(4.39) and Fig. 4.3 to generate a sample function $\{\underline{X}_n\}$, we can obtain $\overline{W}$ and D by setting M large enough, through simulation. The same procedure can be used for IRAR IV-IV_A schemes by incorporating Eqs. (4.37) and (4.21)-(4.22); and similarly for any other IRAR scheme. These simulation procedures have been used to derive the delay-throughput performance curves under IRAR IV-V schemes, which will be presented latter. For an IRAR IV_A scheme, we can compute limiting expectations (3.7b)-(3.7c) and use (3.7a) to analitically obtain an expression for D, in terms of u_o, presented in the following Theorem.

Theorem 4.2. For a multi-access communication channel under IRAR V_A scheme, with $R \geq 1$ and $\lambda < 1$, the limiting average packet delay D is given by

$$D = \overline{W} + 1 + R, \tag{4.40}$$

where

$$\overline{W} = \{E[W(\underline{X}_n, \underline{X}_{n+1})]\} \cdot \{E[N(\underline{X}_n, \underline{X}_{n+1})]\}^{-1} ,$$

and

$$E[W(\underline{X}_n, \underline{X}_{n+1})] = \frac{1}{2}\overline{X^2}(1+\lambda) + \overline{X}\{R\lambda[1-\exp(-\lambda)] + \frac{1}{2}\lambda + R + \frac{1}{2}\}$$

$$+ R\lambda[1-\exp(-\lambda)][\frac{1}{2}(R+1) + u_o R] + \frac{1}{2}u_o\lambda R(R+1) , \tag{4.41}$$

$$E[N(\underline{X}_n, \underline{X}_{n+1})] = \overline{X} + R\lambda\exp(-\lambda) . \tag{4.42}$$

The limiting moments are given by

$$\overline{X} = \lambda(1-\lambda)^{-1}\{R[1-\exp(-\lambda)] + 1 + u_o R\} ,$$

$$\overline{X^2} = (1-\lambda^2)^{-1}\{R\lambda[1+\lambda-\exp(-\lambda)] + R(R-1)\lambda^2[1-\exp(-\lambda)]^2$$

$$+ \lambda + \lambda^2 + \lambda(1+2\lambda)(\overline{X}+u_o R) + \lambda^2 u_o R^2$$

$$+ 2R\lambda^2[1-\exp(-\lambda)][1+\overline{X}+u_o R]\} ,$$

where $u_i = \lim_{n\to\infty} P(X_n=i)$, $\overline{X} = \sum_i iu_i$, $\overline{X^2} = \sum_i i^2 u_i$. For $\lambda \geq 1$, D is unbounded. |

Performance Computations and Comparisons

We first compare the delay-throughput curves for IRAR IV_A and IRAR V_A schemes (given in Theorems 4.1 and 4.2, respectively), and the performance curves for IRAR IV-V schemes (obtained according to Lemma 4.1), with the delay-throughput curves for an IRAR III scheme (obtained by simulation). The resulting curves are shown in Figs. 4.4 - 4.5.

For R = 1, we note that IRAR IV (V) scheme is identical to IRAR $IV_A(V_A)$ scheme. The delay-throughput curves for IRAR III, IV, V procedures, for a channel with R = 1, are shown in Fig. 4.4. We note that IRAR IV (or IV_A) and IRAR III exhibit essentially identical delay-throughput performance curves. IRAR V (or V_A) scheme exhibits a performance curve which is almost identical to those obtained under IRAR III-IV scheme when $\lambda \geq 0.4$, while being somewhat higher at lower network throughput values (as expected, since an initial period starts with any RA slot for IRAR V). We thus conclude that the analytical expressions for the channel delay-throughput curve, when R = 1, under an IRAR IV_A (IV) scheme, essentially yield those resulting under an IRAR III scheme, while those under an IRAR V_A (V) scheme yield a very close lower bound.

For a channel with a higher propagation delay value, such as is the case for a satellite channel, setting R = 12, the resulting performance curves under IRAR III, IV, IV_A, V, V_A schemes are shown in Fig. 4.5. We note again that the performance curves under IRAR IV_A - V_A schemes yield very close upper-bounds to the performance curve under IRAR III scheme, and present almost perfect fit for medium and high network throughput values. The performance curves under IRAR IV - V schemes are noted to yield only somewhat higher packet delay values. As expected, schemes IV and V(IV_A and V_A) exhibit very close performance curves, except that at low network throughput values IRAR V yields somewhat higher packet delay values.

For a channel with a low propagation delay, R = 1, the delay-throughput curves under IRAR I, II, III schemes, as well as under the ARDA I reservation scheme, are shown in Fig. 4.6. We note that the IRAR III scheme yields the lowest average packet delays, over the whole range of network throughput values. For low traffic intensity values, $\lambda < 0.4$, IRAR I, II, III schemes yield the same packet delay values. The latter are lower than those resulting when the pure reservation ARDA I scheme is employed. For high network throughput values, $\lambda > 0.4$, ARDA I and IRAR III schemes yield packet delays values considerably lower than those obtained under IRAR I-II scheme; with the delay values resulting when an IRAR III scheme is employed being uniformly the lowest.

For a channel with a high propagation delay value, $R = 12$, the delay-throughput curves of the various IRAR, reservation and RA schemes are shown in Fig. 4.7. For low network throughput values, $\lambda < 0.3$, IRAR I, II, III schemes yield the same low packet delay values (varying from $D = R + 1 = 13$ slots at $\lambda = 0$ to $D \cong 20$ slots at $\lambda = 0.3$). The same packet delay values, over $0 \leq \lambda < 0.3$, are also obtained by a (stabilized) slotted pure random-access scheme (such as slotted ALOHA or GRA controlled schemes, denoted as RA in Fig. 4.7). The random-access scheme yields, however, packed delay values which rapidly become unbounded as λ approaches a value of $e^{-1} \cong 0.368$ packets/slot. It is interesting to note the IRAR schemes yield as low packet delay values as those obtained under a pure RA scheme for $\lambda \leq 0.3$ (and lower for $\lambda > 0.3$), while no RA retransmissions are being employed by the IRAR schemes. This is explained by noting that the average number of retransmissions per packet (for RA scheme) is lower than one if and only if $\lambda < e^{-1}$ (see [2]), and is thus considerably lower than 1 over the throughput range within which the RA scheme is effective. Subsequently, if a RA scheme induces a high enough probability of packet retransmission, excessive packet delay will result. Therefore, an IRAR scheme, that uses no packet retransmissions at all, can exhibit packet delays as low as those obtained under a pure RA scheme for $\lambda \leq 0.3$. The IRAR schemes, furthermore, clearly exhibit a much superior delay-throughput performance characteristics at higher network throughput values. (Similar observations can be made for $R = 0, 1$.)

At higher network throughput values, $\lambda > 0.3$, we note the IRAR II scheme to yield packet delays considerably lower than those attained under the IRAR I scheme, and somewhat higher than those achieved when the IRAR III scheme is used.

The delay-throughput performance curves of two pure reservation schemes, ARDA I and DFRAC (See Section I for the description of these schemes), are also shown in Fig. 4.7. At low network throughput-values, $0 \leq \lambda \leq 0.4$, similar packet delay values are attained by these two reservation procedures. The latter values are higher than those obtained under the IRAR and RA schemes (by as much as R + 1 = 13 slots, the delay involved in broadcasting the reservation packet, at $\lambda = 0$). Both IRAR II-III schemes are noted to yield delay curves which are lower than the ARDA I delay curve over the whole throughput range. The delay-throughput curve under the DFRAC scheme is above the IRAR III delay curve for $0 \leq \lambda \leq 0.55$; however, for higher network throughput values, $\lambda > 0.55$, the DFRAC scheme yields lower packet delays than those obtained under the IRAR III scheme.

We thus conclude, that to obtain the lowest delay-throughput curve over the whole throughput range, for high values of R, we need to employ a scheme that uses an IRAR III procedure for $0 \leq \lambda \leq 0.6$, and then causes its protocol to transition into that of an ARDA I (or ARDA II, see [1]) pure reservation procedure. This integrated scheme will induce a performance curve following closely the lower envelope of the IRAR III - ARDA I curves. The above transition can be attained, when long-term traffic intensity fluctuations are considered, by incorporating an appropriate estimate of the underlying present network traffic intensity value λ. For short-period fluctuations, a feedback scheme, (such as the ARDA II scheme in [1], incorporating observed reserving packet queue sizes, can be employed to automatically adapt to λ and cause the transition of the protocol of the access-control discipline between the IRAR mode and the pure reservation mode. Such a feedback scheme has been in fact implemented and observed to yield a performance curve following closely the lower envelope of the IRAR III - ARDA I curves.

In Fig. 4.8 we present curves for probabilities u_o, u_1, for R = 1, 12, under IRAR IV_A - V_A schemes. We note that u_o represents the probability of a period with no reserved transmissions, and is used in Theorems 4.1-4.2 to yield explicit analytical solutions for the average packet delays under IRAR IV_A - V_A schemes. Probability $(1-u_o)$ is also a measure of channel utilization.

V. Conclusions

We have presented and studied Integrated Random-Access Reservation (IRAR) schemes for multi-access communication channels. Under an IRAR scheme, newly arrived packets can be designated for reserved or random-access (RA) transmissions. In the latter case, if a collission occurs, each colliding packet is assigned for transmission by reservation, rather than attempt another RA transmission.

The performance of the IRAR schemes under consideration is expressed in terms of the associated channel delay-throughput curves. Analytical expressions have been derived for the limiting average packet delay, under various IRAR schemes, in terms of the steady-state probabilities that an underlying embedded Markov chain is in state 0 (and, sometimes, also in state 1). Single packet messages are considered. We show that for low channel propagation delay values (when considering, for example, terrestrial radio or line networks), an IRAR scheme (IRAR III) yields excellent delay-throughput performance characteristics, over the whole range of network throughput values. The IRAR III performance curve is then also shown to be uniformly superior to that obtained by a corresponding (ARDA I) pure reservation scheme, as well as a corresponding slotted random-access scheme.

For a channel with longer propagation delays (setting R=12), such as a satellite communication channel, we show that the IRAR III scheme yields an excellent delay-throughput curve for network traffic intensity (throughput) values of $0 \le \lambda \le 0.6$. For $0 \le \lambda \le 0.3$, the latter IRAR curve is identical to that obtained when a slotted pure (controlled) random-access procedure is used. The random-access performance curve involves, however, a rapidly increasing packet delay value as $\lambda \uparrow e^{-1} \cong 0.368$. The IRAR performance curve is also superior, within $\lambda < 0.6$, to that obtained when an excellent (considering high propagation delay values) pure reservation scheme (DFRAC) is used.

The latter reservation scheme yields, however, lower average packet delays for higher ($0.6 \leq \lambda < 1$) network throughput values. Therefore, for channels with long propagation delays, low packet delays are obtained over the whole network throughput range, when we employ an access-control scheme that adopts an IRAR protocol for low and medium throughput values, and switches into an efficient pure reservation protocol (as that of a DFRAC or ARDA II schemes) at higher throughput values.

In many actual multi-access communication channel situations, different classes of messages can be distinguished, in accordance with their statistical characteristics and service requirements. An efficient access-control scheme, integrating both random-access and reservation procedures, will then be employed. The basic schemes and the associated analytical techniques presented in this paper, would then provide the necessary basic access-control elements involved in the analysis and design of many such integrated schemes.

References

1. I. Rubin, "Reservation Schemes for Dynamic Packet Access-Control of Multi-Access Communication Channels," Technical Report, UCLA-ENG-7712, School of Engineering and Applied Science, University of California, Los Angeles, January 1977. See also Proceedings of the 1977 Conference on Information Sciences and Systems, The Johns Hopkins University, March, 1977.

2. I. Rubin, "Group Random-Access Disciplines for Multi-Access Broadcast Channels," Technical Report, UCLA-ENG-7745, School of Engineering and Applied Science, University of California, Los Angeles, June 1977.

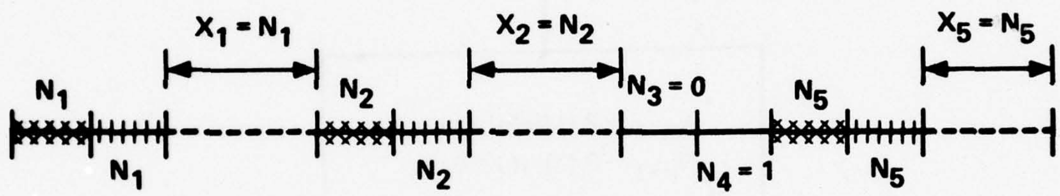

Fig. 3.1. A Sample Function of the Channel State Process Under IRAR I Scheme, When R = 0. A Reservation Slot is Denoted as ┼┼┼┼ , a Slot with Collissions as XXXX , and a Service Period as – – – – .

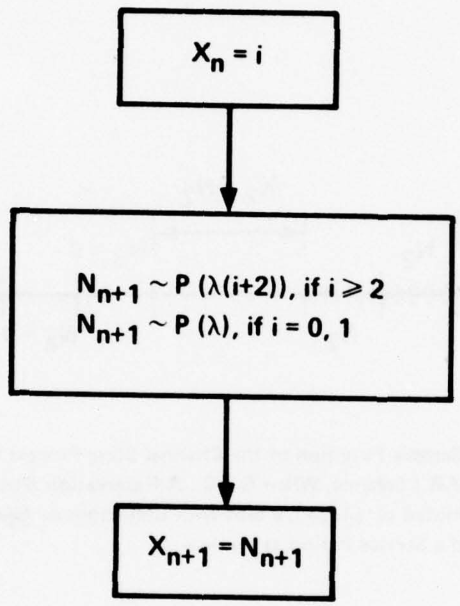

Fig. 3.2. Flow Diagram for the Channel Markov State Sequence Under IRAR I Scheme, When R = 0. P (λ) Denotes a Poisson Distribution with Mean λ .

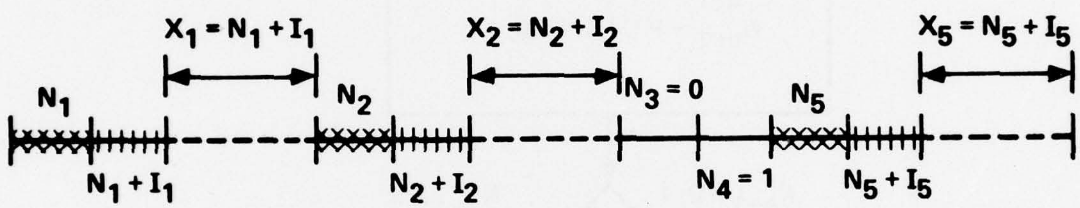

Fig. 3.3. A Sample Function of the Channel State Process Under IRAR II Scheme, When R = 0.

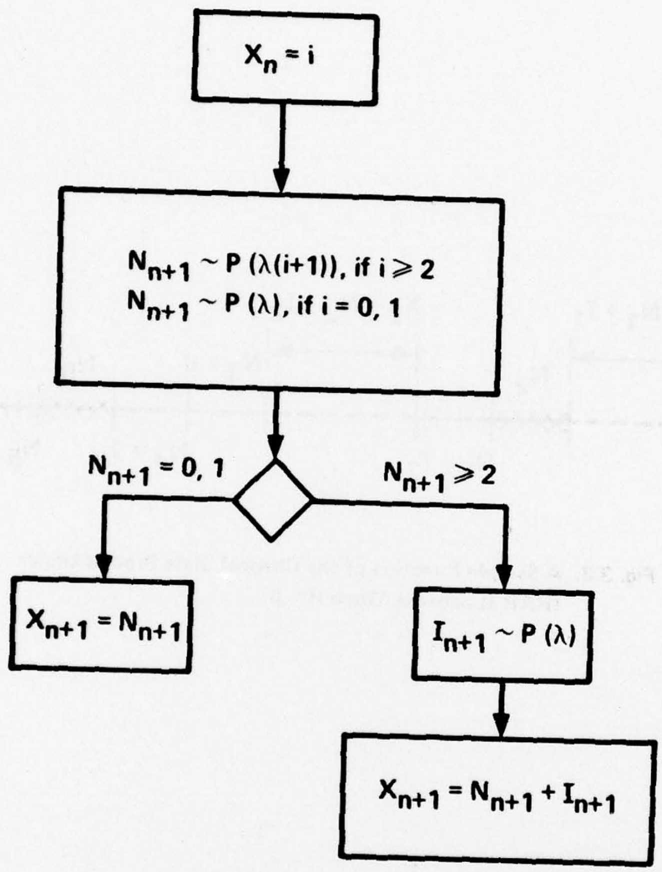

Fig. 3.4. Flow Diagram for the Channel Markov State Sequence Under IRAR II Scheme, When R = 0.

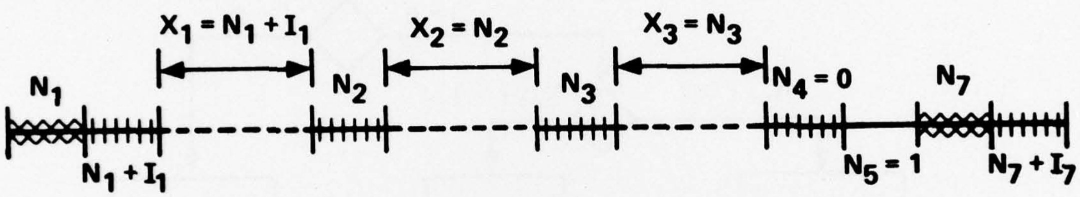

Fig. 3.5. A Sample Function of the Channel State Process Under IRAR III Scheme, With R = 0.

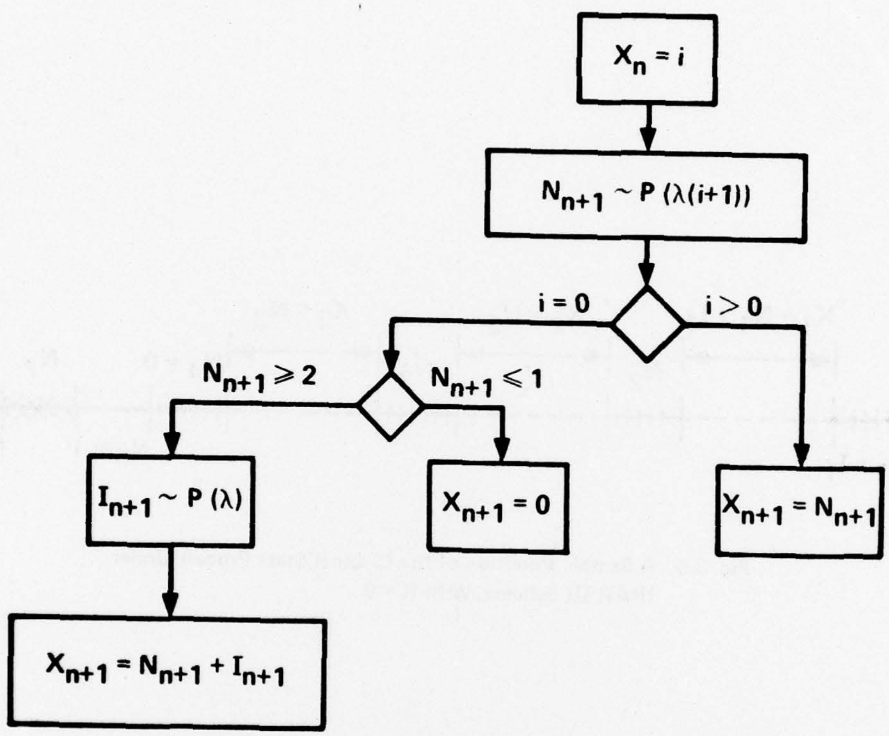

Fig. 3.6. Flow Diagram for the Channel Markov State Sequence Under IRAR III Scheme, When R = 0.

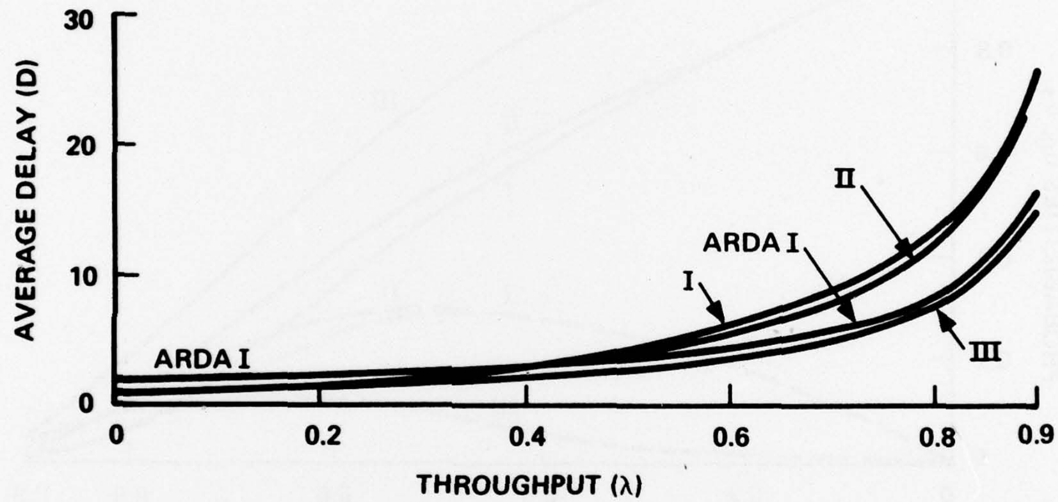

Fig. 3.7. Delay-Throughput Curves for a Channel with R = 0, Under an ARDA I Scheme and IRAR I, II, III Schemes.

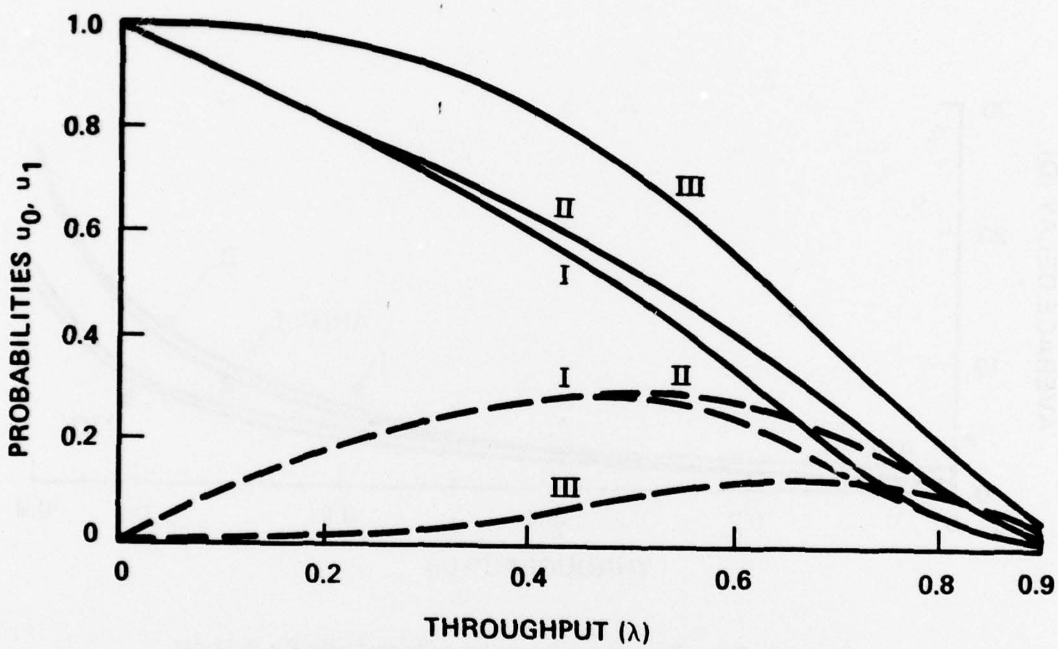

Fig. 3.8. Curves for Probabilities u_o (——) and u_1 (— —) vs λ, for a Channel with R = 0 Under IRAR I, II, III Schemes.

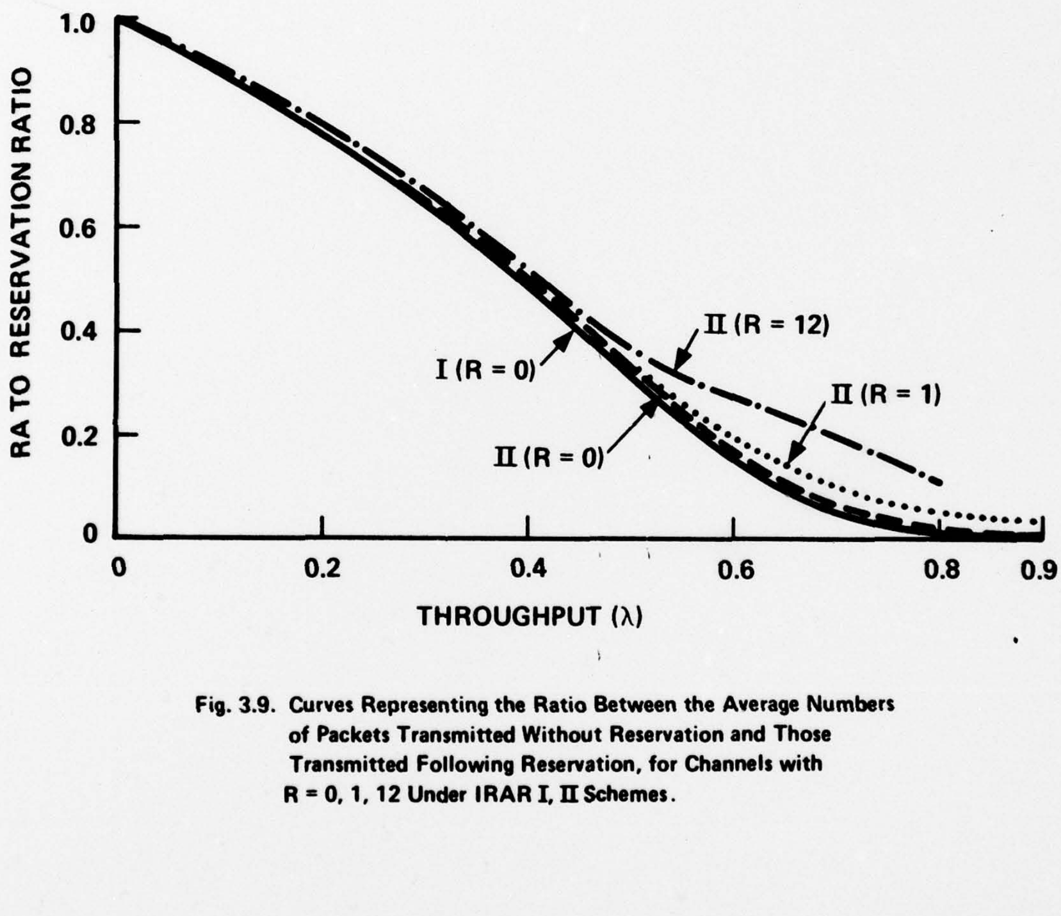

Fig. 3.9. Curves Representing the Ratio Between the Average Numbers of Packets Transmitted Without Reservation and Those Transmitted Following Reservation, for Channels with R = 0, 1, 12 Under IRAR I, II Schemes.

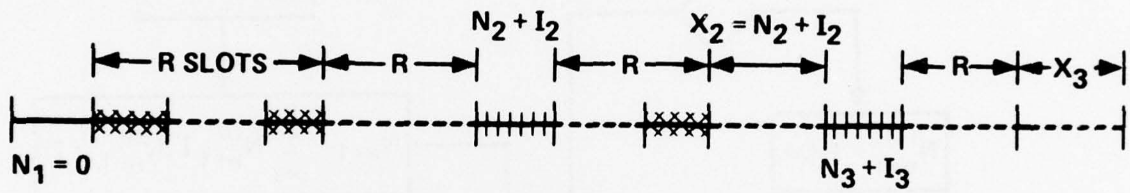

Fig. 4.1 A Sample Function of the Channel State Process Under IRAR IV Scheme, with $R \geqslant 1$.

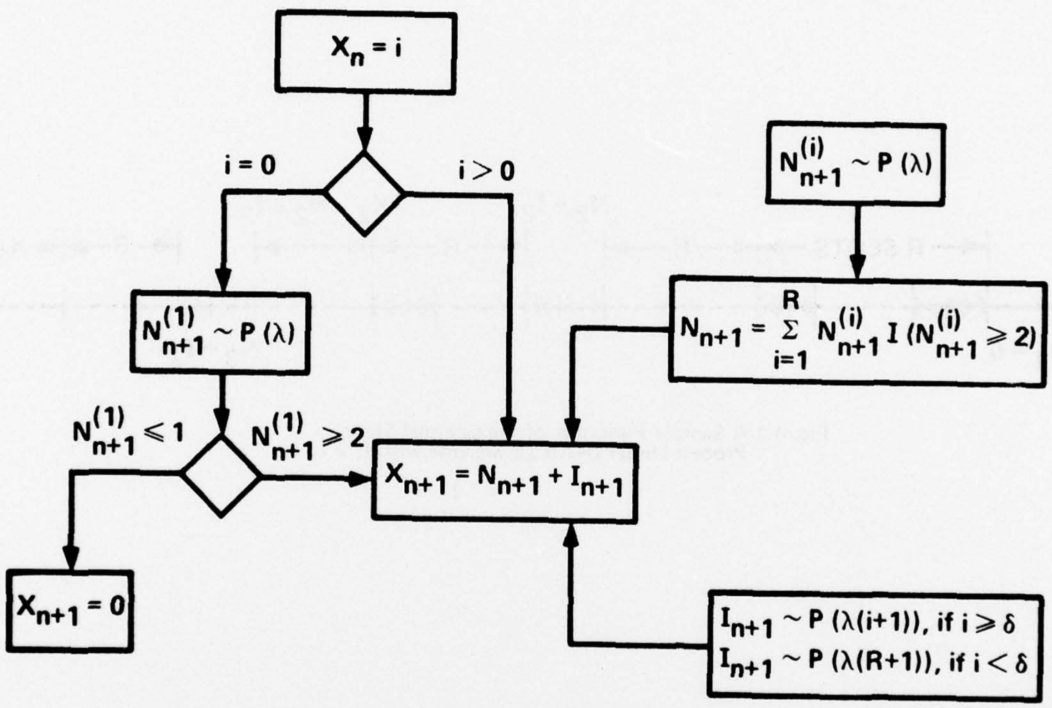

Fig. 4.2. Flow Diagram for the Channel Markov State Sequence Under IRAR IV Schemes, with $R \geqslant 1$. For IRAR IV_A We Set $\delta = 1$, While for IRAR IV We Have $\delta = R$.

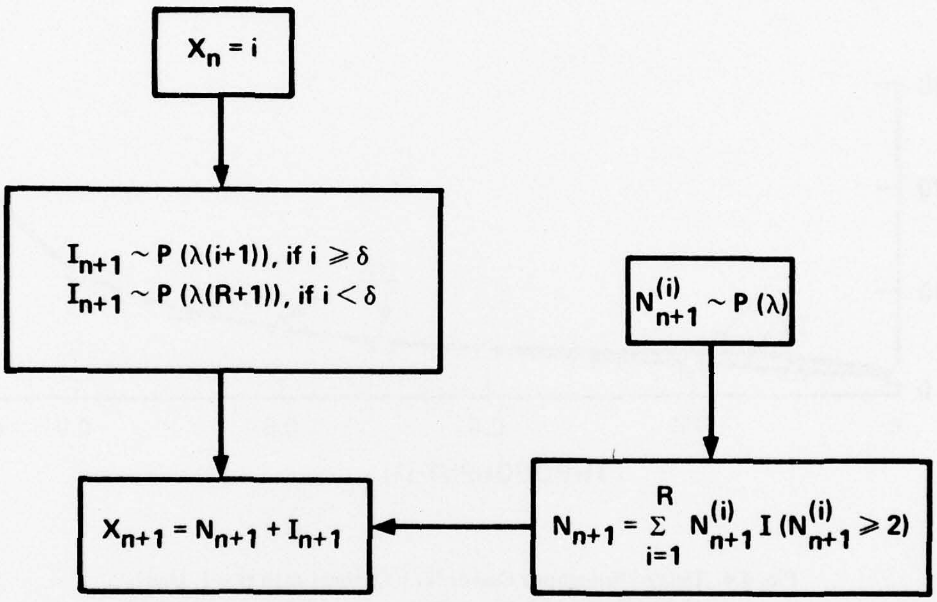

Fig. 4.3. Flow Diagram for the Channel Markov State Sequence Under IRAR V Schemes, with $R \geq 1$. For IRAR V_A We Set $\delta = 1$, While for IRAR V We Have $\delta = R$.

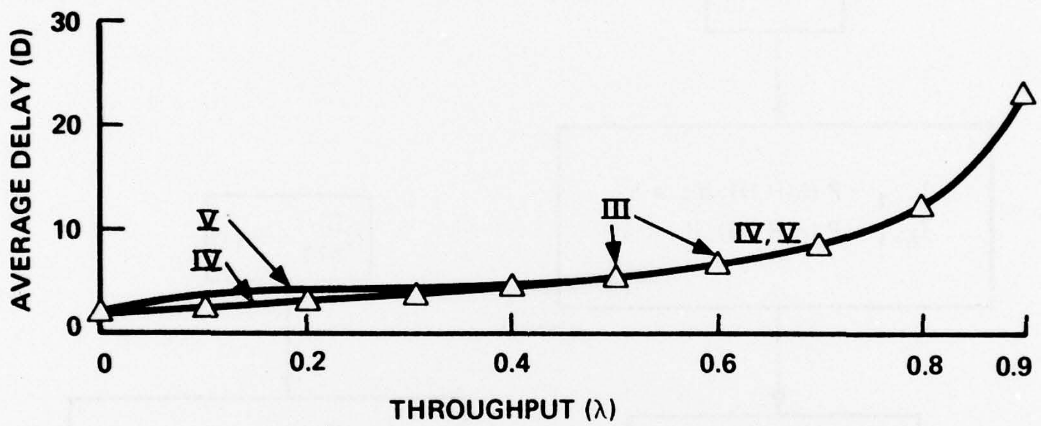

Fig. 4.4. Delay-Throughput Curves for a Channel with R = 1, Under IRAR III, IV, V Schemes. Performance Points for IRAR III are Indicated by △ .

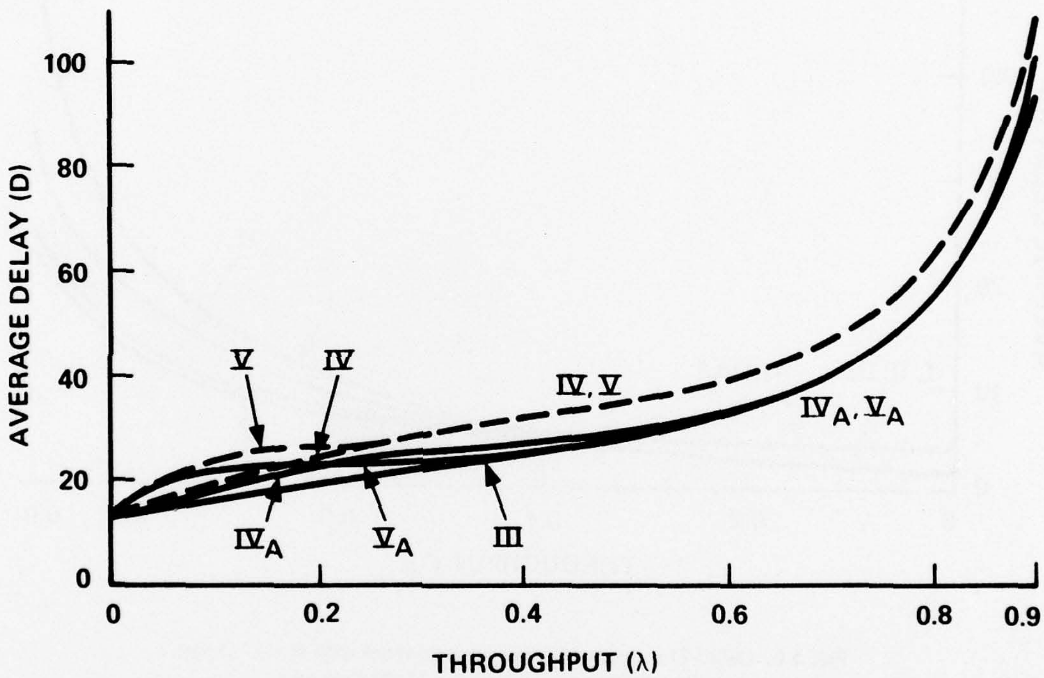

Fig. 4.5. Delay-Throughput Curves for a Channel with R = 12, Under IRAR III, IV, V, IV_A, V_A Schemes.

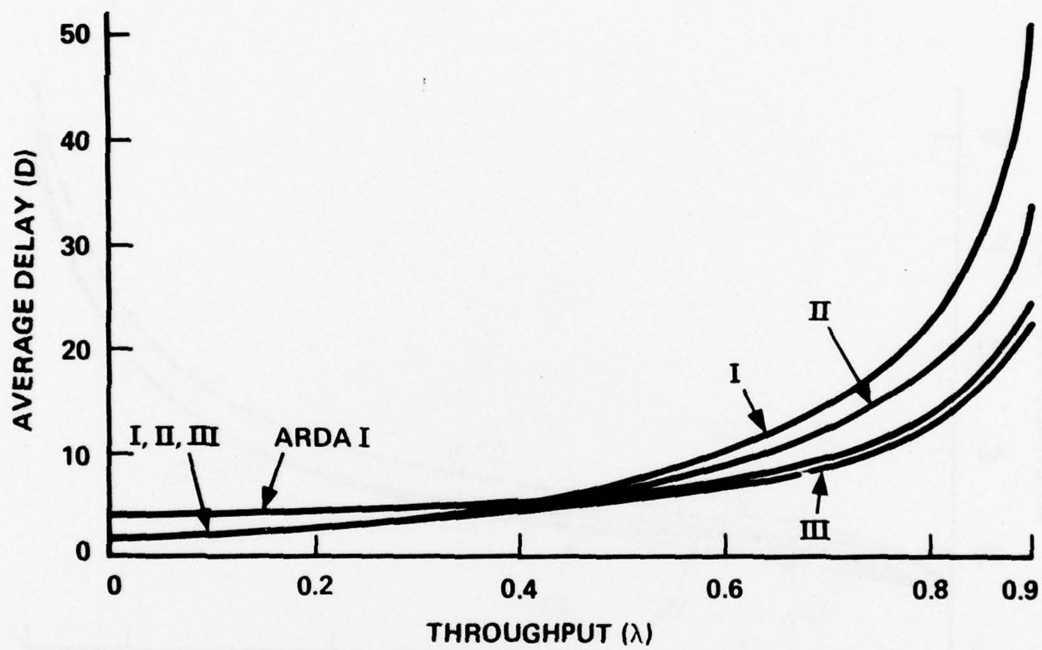

Fig. 4.6. Delay-Throughput Curves for a Channel with R = 1, Under an ARDA I Scheme and IRAR I, II, III Schemes

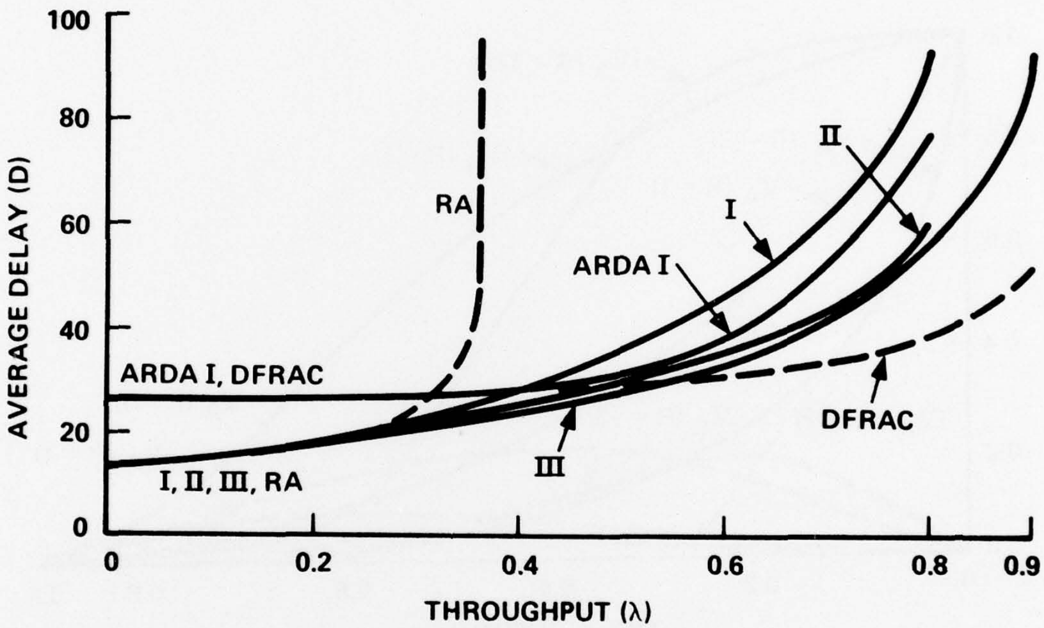

Fig. 4.7. Delay-Throughput Curves for a Channel with R = 12, Under a Slotted Pure Random-Access Scheme (RA), ARDA I Scheme, DFRAC Scheme, and IRAR I, II, III Schemes.

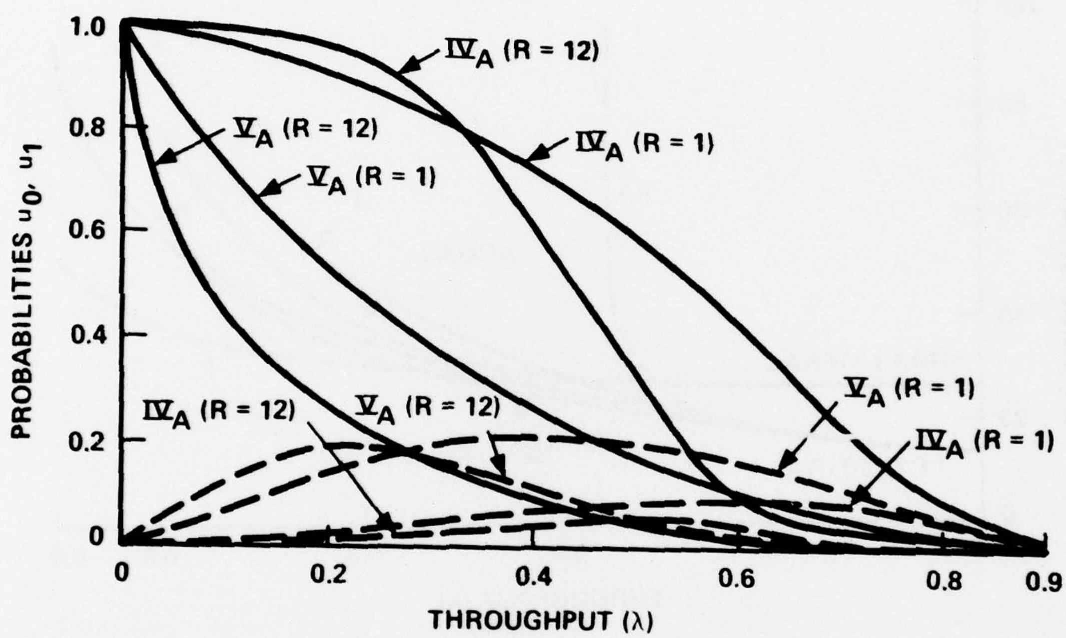

Fig. 4.8. Probabilities u_0 (——) and u_1 (— — —) vs λ, for Channels with R = 1 and R = 12 Under IRAR $\overline{IV}_A$, $\overline{V}_A$ Schemes.